FACULTY OF SCIENCES

Development of seafloor mapping strategies supporting integrated marine management

Application of seafloor backscatter by multibeam echosounders

Giacomo Montereale Gavazzi

Supervisor: Prof. Dr. Vera Van Lancker
Co-supervisors: Dr. Marc Roche, Dr. Xavier Lurton

Submitted to the Faculty of Science of Ghent University in fulfilment of the requirements for the degree of Doctor of Science: Geology

Academic year: 2018 – 2019

GHENT UNIVERSITY

Development of seafloor mapping strategies supporting integrated marine management

Application of seafloor backscatter by multibeam echosounders

Giacomo Ottaviano Andrea Montereale Gavazzi

2019

Submitted for the degree of

Doctor of Science: Geology

Supervisor: Prof. Dr. Vera Van Lancker
Co-supervisors: Dr. Marc Roche, Dr. Xavier Lurton

Members of the examination committee

Prof. Dr. Stephen Louwye (Ghent University, Belgium) Chair
Prof. Dr. Marc De Batist (Ghent University, Belgium) Secretary

Prof. Dr. Steven Degraer (Ghent University; Royal Belgian Institute of Natural Sciences, Belgium)
Prof. Dr. Karline Soetaert (Ghent University; Royal Netherlands Institute for Sea Research, Netherlands)
Dr. Thaiënne van Dijk (University of Illinois Urbana-Champaign, United States of America; Deltares, Netherlands)
Dr. Fantina Madricardo (Istituto di Scienze Marine, Consiglio Nazionale delle Ricerche, Italy)

Prof. Dr. Vera Van Lancker (Ghent University; Royal Belgian Institute of Natural Sciences, Belgium) Supervisor
Dr. Xavier Lurton (Institut Français de Recherche pour l'Exploitation de la Mer, Laboratory of Underwater Acoustics, France) Co-promoter
Dr. Marc Roche (Federal Public Service Economy, Self-Employed and Energy, Continental Shelf Service, Belgium) Co-promoter

Giacomo Montereale Gavazzi received a grant from the Royal Belgian Institute of Natural Sciences, with financial support of the Belgian Scientific Policy Office under the Belgian Action Plan Through Interdisciplinary Networks (BELSPO BRAIN-BE; Contract Grant Nr. BR/143/A2/INDI67).

This doctoral thesis was promoted by Ghent University, Renard Centre of Marine Geology (RCMG) and was primarily carried out at the Royal Belgian Institute of Natural Sciences (RBINS), Operational Directorate Natural Environments. Research was done in close cooperation with the Federal Public Service Economy, of Belgium, Continental Shelf Service (CSS-FPS) and the Institut Français de Recherche pour l'Exploitation de la Mer, Laboratory of Underwater Acoustics (IFREMER – Centre of Brest, Plouzané, France).

To refer to this thesis: Montereale Gavazzi, G., O., A., 2019. Development of seafloor mapping strategies supporting integrated marine management: application of seafloor backscatter by multibeam echosounders. PhD Thesis, Ghent University, Ghent, Belgium.

Research highlights

- **Seafloor mapping is the fundamental and indispensable basis of marine environmental monitoring.** Multibeam echosounding occupies a central role in setting up seafloor mapping and monitoring strategies, maximising survey time and reducing costs. The baseline survey effort towards the implementation of the European Marine Strategy Framework Directive (MSFD, 2008/56/EC) in Belgian waters was successfully planned, acquired and compiled, advancing the long-term, site-specific and regional monitoring of seafloor integrity (MSFD Descriptor 6).

- **Automated integration of multibeam and ground-truth data allows production of accurate and widely applicable habitat maps.** State-of-the-art Acoustic Seafloor Classification and data-integration routines allowed the production of accurate, repeatable and spatially-explicit models of the seafloor nature, maximising the information content achievable from multibeam bathymetry, backscatter and their derivatives. The latter are fundamental proxies of the substrate type, a keystone building block of benthic habitats, allowing obtaining information at scales relevant for ecological management.

- **Knowledge of environmental variability is critical to deal with the dynamic operational environment, as well as to interpret static and serial MBES backscatter datasets.** Dedicated field experiments provided a baseline to quantify and discern between the intrinsic and unwanted types of environmental variability that influence the multibeam backscatter measurements: knowledge necessary to advance the interpretation of serial multibeam datasets.

- **Acoustic change detection is a first critical step to assess and understand the evolution in environmental status of the seafloor.** Methodologies that allow quantifying the signals of seafloor change are needed. Deriving categorical patterns and trends of persistence and from-to transitions from multibeam acoustic imagery is critical to ultimately decipher naturally- from anthropogenically-induced sediment dynamics and is pivotal in the design of monitoring surveys.

Table of contents

List of Figures

List of Tables

List of Abbreviations

ABS	Acoustic Backscatter Sensor
ACD	Acoustic Change Detection
ADV	Acoustic Doppler Velocimeter
APL	Acoustics Physics Laboratory
AR	Angular Response
ARA	Angular Range Analysis
ASC	Acoustic Seafloor Classification
BCS	Belgian Continental Shelf
BHM	Benthic Habitat Mapping
BPNS	Belgian Part of the North Sea
BS	Backscatter Strength
BSWG	Backscatter Working Group
BT	Bottom Trawling
CBI	Compensated Backscatter Imagery
CSS	Continental Shelf Service (BE)
CTD	Conductivity Temperature Depth
CU	Central Unit
DD	Dredging and Dumping
DGPS	Differential Global Positioning System
DP	Dynamic Positioning
EBM	Ecosystem Based Management
EEZ	Exclusive Economic Zone
FF	Far Field
FPS	Federal Public Service (BE)
GEOHAB	Geological and Biological Habitat Mapping
GES	Good Environmental Status
GIS	Geographical Information System
GLCM	Grey Level Co-occurrence Matrices
GSAB	Generic Seafloor Acoustic Backscatter
HCMS	High Concentration Mud Suspension
HCS	Habitat Classification Scheme
HV	Hydrographic Vessel
IFREMER	Institut Français de Recherche pour l'Exploitation de la Mer
IMU	Inertial Motion Unit
LAT	Lowest Astronomical Tide
LIDAR	Light Detection and Ranging
LISST	Laser in Situ Transmissometer
LULC	Land Use Land Cover
MAE	Marine Aggregate Extraction
MBES	Multibeam Echosounder System
MIDAS	Marine Information and Data Acquisition System
MLC	Maximum Likelihood Classifier
MSP	Maritime Spatial Planning
NF	Near Field
NOC	National Oceanography Centre (UK)

OBS	Optical Backscatter Sensor
ODAS	On-Board Data Acquisition System
PDF	Probability Density Function
QF	Quality Factor
RADAR	Radio Detection and Ranging
RCMG	Renard Centre of Marine Geology – University Ghent
RF	Random Forest
ROI	Region of Interest
ROV	Remotely Operated Vehicle
RTK	Real Time Kinematic
RV	Research Vessel
SAR	Synthetic Aperture Radar
SAV	Submerged Aquatic Vegetation
SBES	Single Beam Echosounder System
SEACOP	Surveying for Environmental Assessments Community of Practice
SNR	Sound to Noise Ratio
SONAR	Sound Navigation and Ranging
SPM	Suspended Particulate Matter
SSS	Side Scan Sonar
SVP	Sound Velocity Profiler
TL	Transmission Loss
TWTT	Two-Way Travel Time
UAV	Underwater Autonomous Vehicle
UTM	Universal Transverse Mercator
VARIMONIT	A GEOHAB-BSWG piloted subgroup on Variability and Monitoring
VLIZ	Flanders Marine Institute (Vlaams Instituut voor de Zee)
WCB	Water Column Backscatter
WE	Wind Energy
WGS	World Geodetic System

Acknowledgments

There are several individuals I am deeply grateful to and would like to duly and warmly thank for having me inspired and supported throughout these years and my life in general.

First and foremost, I would like to thank my parents. My mother Francesca for having me made, shaped and taught to contemplate and enjoy life and nature in the best manner I could imagine and for having me encouraged with love to pursue my interests, always. Equally, my father Prof. Giuseppe, for truly inspiring me as a father and as a nature lover and insightful observer, and for supporting me with love from the very first moment I met him.

I am enormously grateful to my supervisor Prof. Vera Van Lancker and co-promoters Dr. Marc Roche and Dr. Xavier Lurton for constructively supporting my ideas and my desk- and field-work constantly and with enthusiasm, and for sharing with me their skills and knowledge.

I would like to thank all colleagues from the SUMO group of the Operational Directorate of Nature of the Royal Belgian Institute of Natural Sciences, in particular those which took part in the field-work. I deeply acknowledge my close collaborators Dr. Nathan Terseleer, Frederic Francken and Lars Kint for sharing with me expertise and good mood. I also deeply thank Dr. Michael Fettweis and Yi Ming Gan for truly enriching my days at RBINS. I am grateful to Reinhilde Van den Branden that often helped me onboard and has truly supported several oceanographic campaigns.

This work would have not been possible without the managers, crews, commanders and captains of the research vessels *Belgica* and *Simon Stevin*, whom I thank deeply for their professional field support and naval-related conversations from which I could learn about interesting and important naval-related knowledge.

Michiel T' Jampens from VLIZ is acknowledged for providing technical expertise and good will in a side project improving underwater videographic sampling in these murky waters.

Dr. Inka Meyer and Professor Sebastien Bertrand are duly acknowledged for their help and advice on the use of sediment analysis instrumentation at the RCMG facilities.

Samuel Deleu of Flanders Hydrography and Koen Degrendele of the Continental Shelf Service of the Federal Public Service Economy Self Employed and Energy of Belgium are deeply thanked for their help with tidal datasets, crucial for the hydroacoustic data processing, and for generally very actively collaborating during various stages of the thesis.

All the researchers I met during my visits in Brest visiting Doctor Xavier Lurton at the Laboratory of Acoustics of IFREMER have been truly inspiring individuals and I would like to thank them warmly, in particular Dr. Jean-Marie Augustin and Dr. Fabrice Lecornu. Equally, the expertise and support I received from Dr. Marc Roche (FPS-CSS), have enabled me to gain confidence in this largely transdisciplinary project.

I thank from the deep of my heart doctors Fantina Madricardo and Federica Foglini (ISMAR-CNR, IT) and Aleksandra Kruss (ISMAR-Norbit) for having opened the most important doors of my academic life and being truly inspirational researchers with whom I wish to continue collaborating *ad vitam*. Equally I thank Prof. Daniel Ierodiaconou from Deakin University Australia for having me enormously inspired and for having me always encouraged.

I want to thank two PhD students I met at conferences during this time of my life with whom I will share marvellous memories forever: Daphnie Galvez and Timo Gaida.

I also would like to thank my most recent scientific collaborators, Ting Zhao and Prof. Dr. Aleksandra Pizurica from the Department of Telecommunication and Information Processing (TELIN-Ghent University).

I sincerely thank the peer-reviewers from the scientific journals whose comments helped me improve all the manuscripts I published and that I present or have co-author and mention in this work.

I want to enormously and deeply thank Chantal Nassef Magdy for ever supporting me and for being the best marine treasure hunting partner.

Special gratitude goes to my friends! Ah, my dearest friends, what would the world be without you? A barren world. Special thanks to my dearest friends Dr. Chief Astrophysicist Marco Rocchetto, Marcel Debczyński, Egon Ustino, Enrico Grimani, Andrea Santoro and Federico Infelise, Alessandro Pietromartire and Maria Laura Carozza: ciao fioiii! Bea fioiii! Also, I would like to deeply thank Christophe, friend and owner of the *Repaire du Sommelier.*

Finally, this work would have certainly not been possible without the crucial and constant presence of Radiotelevisione Italiana RAI Radio 3 classical music broadcasts and the music produced by Goldie, Massive Attack and the 4-Rooms family.

This work is dedicated to the memory of Kala': may the most beautiful pheasants, pigeons and cormorants be with you always.

Foreword

Chapters 4 and 5 have been published in international peer-reviewed journals and Chapter 3 is in preparation. Due to this style of presentation, I would like to apologise to the reader for the moderate redundancies among the introductory sections of the chapters.

All cartographic figures are projected in UTM zone 31N-WGS 84.

Samenvatting

Een wereldwijde synergie van alomtegenwoordige antropogene druk bedreigt het mariene ecosysteem. Dit is gerelateerd aan toenemende zeebodemexploitaties, alsook grootschalige kust- en offshore-infrastructuurontwikkelingen (Halpern et al., 2008, 2015), gestuwd door de moderne Westerse economische golf: Een oceanische "goud en energiekoorts".

Meer dan ooit is er behoefte aan een verantwoord beheer van mariene hulpbronnen (Pickrill en Kostylev, 2007), idealiter gebaseerd op zeebodemkartering die toelaat de mate van beïnvloeding te onderzoeken, te evalueren en in de tijd op te volgen (Tekman et al., 2017; Woodall et al., 2014). Intussen is de karteringstechnologie (i.e., multibeam echosounders; Lurton, 2010) dusdanig ontwikkeld dat deze een revolutie teweegbracht in ons vermogen om de zeebodem akoestisch in beeld te brengen. Geavanceerde metingen van diepte en terugverstrooiingswaarden van het akoestisch signaal ('backscatter') zijn nu in staat om de geometrie van de zeebodem en haar natuur in detail (fijnmazig) en op continue ruimtelijke schalen (meso- tot grootschalige schaal) te kwantificeren, van cruciaal belang voor het voorspellen van zeebodembiodiversiteit, maar ook ter ondersteuning van tal van andere toepassingen. Toch zijn er nog heel wat uitdagingen. Op wereldschaal worden ambities geformuleerd om de hele onderwaterwereld aan een hoge resolutie in kaart te brengen (Seabed2030, Mayer et al., 2018); op de meer regionale schaal is er nood aan gekalibreerde multi-parameter datasets om, in een cyclisch proces, de milieutoestand van het mariene ecosysteem te beoordelen (i.e., de Europese Kaderrichtlijn Mariene Strategie (KRMS, 2008/56/EG); bijv. Madricardo et al., 2017).

In deze context is de kwaliteit, efficiëntie en herhaalbaarheid van multibeamopnames, en van het afleiden van datatypes, heel belangrijk. Kalibratie met bemonsteringen en visuele observaties is cruciaal, waarbij data-integratie met behulp van innovatieve automatische classificeringsroutines (i.e., akoestische zeebodemclassificatie; Anderson et al., 2007, 2008) wordt vooropgesteld. De noodzakelijke transdisciplinariteit in dit onderzoeksgebied weerspiegelt onze erkenning van de complexiteit van het mariene ecosysteem dat verder dient ontdekt en onderzocht te worden, in het bijzonder voor de opvolging van mogelijke negatieve gevolgen van de toenemende menselijke druk. De nood aan ontwikkeling van akoestische classificatie- en veranderingsdetectie-strategieën weerspiegelt evenzeer de noodzaak om snelle, nauwkeurige en budgetefficiënte oplossingen aan te reiken ter ondersteuning van wetenschappelijk advies voor een beheer van een omgeving waarvan we de milieuwaarde nog maar recent erkennen.

Dit geldt in het bijzonder voor het Belgische deel van de Noordzee (BDNZ), één van wereld's dichts bevaren gebieden en met een sterke concentratie van menselijke drukken (Douvere et al., 2007). Toch zijn regionale zeebodembedekkende kaarten die een opvolging van de milieutoestand toelaten beperkt. Dit is gerelateerd aan de

veelheid aan antropogene activiteiten die continue en wijdverspreide metingen bemoeilijkt, alsook aan het ontbreken van een gestandardiseerd, en nationaal gecoördineerd, zeebodemkarteringsprogramma, ondanks de wettelijke verplichtingen die Europa oplegt inzake milieuopvolging.

In deze sociaal-juridische context biedt het BDNZ het optimale operationele laboratorium om de principes van akoestische zeebodemkartering te onderzoeken en te testen met het oog op de ontwikkeling van tijdsefficiënte opvolgingsmethodologieën van de milieutoestand, temeer deze nodig zijn om de effecten van menselijke activiteiten in Belgische en andere wateren te meten, en op te volgen. Dit onderzoek kadert dan ook in de stroom van kennis die de disciplines van onderwaterteledetectie en mariene ecologie *s.l.* trachten op te bouwen in een complex milieu dat relatief moeilijk in de ruimte en tijd te beschrijven is.

Het algemene doel van dit doctoraatsonderzoek was om verschillende toepassingen van akoestische meettechnologie nader te bestuderen zoals kartering, maar evenzeer classificatie van de zeebodem en detectie van veranderingen van een milieutoestand. Een initieel doel was het definiëren van een basisinspanning van herhaalbare metingen en benaderingen waarop een aantal hedendaagse data-integratieroutines konden worden toegepast, en waarbij de nauwkeurigheid en herhaalbaarheid kon getest worden op complexe multivariate datastructuren. Het was belangrijk nauwkeurige en gedetailleerde zeebodemmodellen te ontwikkelen, zowel vanuit een statisch als tijdsdynamisch standpunt, en aldus te bouwen aan een verbetering van de Belgische (en Europese) KRMS-monitoring. De betrouwbaarheid en herhaalbaarheid van multibeamterugverstrooiingswaarden stonden centraal, aangezien deze gegevens erkend worden als unieke en fundamentele geofysische informatie die de kartering van benthische substraten en habitats onderbouwt. In deze context was de kwantificering van de omgevingsvariabiliteit van cruciaal belang omdat verschillende bronnen van variantie de hydro-akoestische metingen kunnen beïnvloeden en gevolgen kunnen hebben voor de interpretatie van kaartproducten, momentaan en in de tijd. Tenslotte dienden akoestische benaderingen voor het opsporen van veranderingen in de zeebodem onderzocht te worden in functie van veranderingen in een milieutoestand, evenals testmethoden om de overheersende signalen van verandering te duiden en deze in een volgende en/of synchrone fase te koppelen aan causale factoren.

Dit proefschrift begint met een introductie van de fundamentele technische en fysische achtergrond van hydro-akoestische metingen, de afgeleide datatypes en hun integratie (hoofdstukken 1a, 1b en 2). Vervolgens worden drie belangrijke onderzoekslijnen uitgewerkt (hoofdstukken 3, 4 en 5), gevolgd door een uitgebreide discussie over de resultaten en uitdagingen die zich in de loop van het onderzoek hebben voorgedaan, alsook de verdere stappen en toekomstmogelijkheden (hoofdstuk 6). Hoofdstuk 7 bevat een algemene conclusie.

In Hoofdstuk 3 worden sediment-akoestische relaties gekwantificeerd die noodzakelijk zijn om, op basis van bemonsteringen en visuele observaties, de akoestische metingen te gebruiken voor het voorspellend in kaart brengen van benthische substraten. Een grote meerbronnige dataset werd geproduceerd op basis van een strategische meet- en verwerkingsstrategie die het mogelijk maakte om een naadloze, continue kaart te bekomen van de multibeamterugverstrooiingswaarden. Dit liet toe benthische substraattypes te karteren wat aan de basis ligt van het voorspellen van het voorkomen van macrobenthische gemeenschappen en hun status. Harde/heterogene substraatgebieden werden in detail gekarakteriseerd gezien hun grote belang voor de identificatie van biodiversiteitshotspots. De terugverstrooiingswaarden en de dieptegegevens werden tevens gebruikt om de voorspellende performantie van twee statistische classificeerders voor de productie van ruimtelijk-expliciete modellen van benthische substraten te testen: een niet gecontroleerde en een gecontroleerde classificatieroutine, respectievelijk 'k-means' partitieve clustering en 'Random Forest Machine Learning' classificatie. Voorts werden testen uitgevoerd naar het onderscheidend vermogen van multibeamverstrooiingswaarden bij 300 kHz aan de hand van twee zeebodemclassificatieschema's , en naar verschillende benaderingen om het aantal clusters (empirisch en/of statistisch) te vinden. Een waaier van nauwkeurigheidsmetriek bevestigde de sterkte en de zwakheden van de geautomatiseerde classificatievoorspellingen op basis van (1) terugverstrooiingswaarden alleen, (2) terugverstrooiingswaarden en diepte, en (3) met inbegrip van een reeks relevante morfometrische en textuur-gerelateerde afgeleiden. Vergelijking en evaluatie bevestigde de effectiviteit van de gecontroleerde multivariate 'Machine Learning' benadering. De resultaten van deze analyses en de voorspellingskracht werden beoordeeld door het implementeren van een grondig protocol van foutenschatting gebaseerd op de 'confusion' matrix. Verder onderzoek is nodig naar de performantie van de beschikbare classificeerders in andere gebieden, alsook naar de toepasbaarheid van de uitgewerkte classificatieroutines in milieus waarvoor andere datastructuren beschikbaar zijn. In alle gevallen is het bij het voorschrijven van classificatieschema's belangrijk zich bewust te zijn van de keuze van het aggregatieniveau van een classificatieschema, aangezien dit kan leiden tot het verlies van belangrijke milieu-informatie.

In Hoofdstuk 4 worden bronnen van variantie gekwantificeerd, veroorzaakt door kortetermijn-, halfdagelijkse getijdenvariabiliteit die multibeamterugverstrooiingswaarden beïnvloeden. Dit kan gevolgen hebben voor de interpretatie van patronen en trends in seriële terugverstrooiingsdatasets en dus ook voor de beoordeling van de natuurlijk en/of antropogeen geïnduceerde variabiliteit van zeebodemsubstraten. Bijzondere aandacht werd besteed aan de hydrologische toestand en de akoestische transmissieverliezen in de waterkolom, gezien het belang om de terugverstrooiingswaarden uitsluitend te kunnen relateren aan het doelwit: de interface tussen water en sediment. Idealiter zouden de resultaten van

echosounders over de hele wereld vergelijkbaar moeten zijn, zodat het samenbrengen van datasets tot één grote geografische dekking zou leiden. Dit is slechts mogelijk door strenge normering van de manier waarop terugverstrooiingswaarden worden bekomen, alsook mits controle van de omgevingsdrift. Dit is van het grootste belang om onderscheid te kunnen maken tussen veranderingen die het gevolg zijn van veranderingen in zeebodemsubstraat, en veranderingen die het gevolg zijn van andere processen en eigenschappen (bijvoorbeeld de status van de waterkolom, de geometrie van de zeebodem, de vaarrichting). Hiertoe werden drie experimenten uitgevoerd in drie verschillende substraattypes met een transect dat herhaaldelijk werd opgemeten tijdens één getijdencyclus (i.e., +/- 13h). Het opbouwen van basisdatasets die toelaten om bronnen van variantie, hun type (d.w.z. intrinsiek of ongewenst) en omvang te kwantificeren, zijn belangrijk om de gevoeligheid van de hydro-akoestische metingen te schatten, alsook om bibliotheken te bouwen met akoestische kenmerken die representatief zijn voor bepaalde zeebodemtypes.

Hoofdstuk 5 richt zich op het innovatieve onderwerp van akoestische veranderingsdetectie met als doel het testen en evalueren van opvolgingsbenaderingen waarbij nieuwe metingen eerst relatief gekalibeerd worden ten opzichte van een stabiel en natuurlijk gebied met gekende terugverstrooiingswaarden (pragmatische oplossing om de herhaalbaarheid van metingen te beheersen). Dergelijke gecorrigeerde tijdsreeksen van terugverstrooiingswaarden maakt het vervolgens mogelijk om pré-, post- en ensemble classificatiemethodologieën te onderzoeken die nodig zijn om veranderingen in een milieutoestand te meten. Hiertoe werden tevens benaderingen vanuit de terrestrische teledetectiegemeenschap voorgesteld: d.w.z. detectie van spatio-temporele trends en kwantificering van de dominante signalen van zeebodemveranderingen, zoals persistentie, winst, verlies en van-tot-overgangen tussen geclassificeerde scènes.

Tenslotte wordt in hoofdstuk 6 de algemene Discussie voorgesteld, waarbij de vooruitgang en beperkingen die in het onderzoeksproces zijn geïdentificeerd worden herhaald en verder worden uitgewerkt. Verdere onderzoeksmogelijkheden worden gepresenteerd en de resultaten worden besproken in de context van de akoestische classificatie van de zeebodem en de detectie van veranderingen ter ondersteuning van de opvolging van een milieutoestand. Het belang wordt herhaald van de stabiliteit en de herhaalbaarheid van de metingen van de terugverstrooiings-waarden en de wijze waarop de controle ervan het mogelijk maakte gegevens in ruimte en tijd samen te voegen en te vergelijken. Een kritische evaluatie van de data-integratie methodiek (classificatie) wordt voorgesteld, waarbij de meest significante problemen worden geïdentificeerd, d.w.z. moeilijkheden bij het vinden van het optimale aantal klassen, de keuze van zeebodemclassificatieschema's en fundamentele fysische eigenschappen die het discriminatiepotentieel van 300 kHz terugverstrooiingswaarden kunnen beïnvloeden. De ecologische waarde van de

meest accurate voorspellingsmodellen wordt bediscussieerd, alsook hun ruimtelijke onzekerheid. Ten slotte worden verdere inzichten gegeven in de uitdagingen waarmee de zeebodemkarterings-gemeenschap wordt geconfronteerd bij het toepassen van de principes van akoestische veranderingsdetectie.

Summary

A global synergy of pervasive anthropogenic pressures threatens marine ecosystems through increasing large-scale coastal and offshore infrastructural developments (Halpern et al., 2008, 2015), collectively triggered by the modern Western-world-driven economical wave: An oceanic “gold and energy rush”.

More than ever a responsible stewardship (Pickrill and Kostylev, 2007) to the use of marine resources is needed requiring seafloor mapping to explore, evaluate and monitor marine areas affected by such pressures (Tekman et al., 2017; Woodall et al., 2014). Meanwhile, mapping technology (i.e., multibeam echosounders, Lurton, 2010) is in place and has revolutionised our ability to image the ocean floor acoustically. Nowadays, state-of-the-art hydroacoustic measurements, bathymetry and backscatter, allow the characterisation of the seafloor geometry and nature in detail (fine-scale) and at continuous spatial scales (meso- to broad-scale) being of major importance in predicting biodiversity and serving other numerous uses. On a global scale, ambitions are set to map the entirety of the submerged world (Seabed2030, Mayer et al., 2018), though at the more regional scale challenges relate to the compilation of calibrated multi-parameter datasets to assess, in a cyclical process, the environmental status of the marine ecosystem (i.e. European Marine Strategy Framework Directive (MSFD, 2008/56/EC); e.g. Madricardo et al., 2017).

In this thesis the development, testing, cross-validation and investigation of the quality, efficiency and repeatability of multibeam technology and its data types was targeted, through calibration against ground-truthing approaches, and integration using innovative automated classification routines (i.e., Acoustic Seafloor Classification; Anderson et al., 2007, 2008)). The needed transdisciplinarity in this research field reflects our recognition of the intricacy of the system we seek to discover and investigate, particularly in respect to the possibly adverse repercussion our manifold activities exert on the marine ecosystem. Development of acoustic seafloor classification and change detection strategies equally reflects the need to provide fast, accurate and inexpensive solutions to scientifically advice the management of an environment we are only beginning to fully acknowledge.

This is particularly the case for the Belgian Part of the North Sea: nested in a maritime area of the world where mankind’s turmoil is dense and frequently distributed over a limited spatial extent (Douvere et al., 2007). Issues inherent to obtain regional full-coverage mapping are linked to the overarching number of anthropogenic activities and, more generally, to the lack of a nationally coordinated seafloor mapping programme despite the legal obligations enacted at the European level that mandate environmental monitoring.

This sociolegal and geographical setting provides the optimal *in situ* operational laboratory to investigate and test principles of acoustic seafloor mapping and set up

baseline environmental status monitoring efforts and methodologies that are imminently required to gauge impact assessments in Belgian waters and farther afield. This doctoral thesis frames into the stream of spurring knowledge that the underwater remote sensing and marine ecology disciplines are seeking to build towards the improvement of our ability to deal with the harsh operational environment within which measurements are made, and our ability to remotely describe the seafloor in space and time.

To that end, the general aim of this PhD research was to investigate various aspects of acoustic seafloor mapping, classification and monitoring (or change detection): it started with setting up a baseline effort of repeatable surveying efforts and approaches, applying and testing the accuracy and repeatability of some of the state-of-the-art data-integration routines that allow an effective treatment of complex multivariate data structures, creating accurate and detailed models of this barely visible hidden realm, both from a static and temporally dynamic point of view, and building towards the improvement of the Belgian (and European) MSFD monitoring approaches. The reliability and repeatability of multibeam backscatter measurements were dealt with particularly, since these data are recognised as a unique and fundamental geophysical information underpinning the mapping of benthic substrates and habitats. In this context quantification of environmental variability was critical since various sources of variance can influence the hydroacoustic measurements and have repercussions on the interpretation and utilisations of static and repeated mapping. Further, acoustic seafloor change-detection approaches, the inevitable evolution of ASC from an otherwise statically perceived seafloor, were investigated in light of environmental monitoring applications, testing approaches that can be used to capture the predominant signals of change and ultimately, in a following and/or synchronous phase, link those to causal factors.

This thesis starts with an introduction of the fundamental technical and physical background of hydroacoustic measurements, as well as the data types and their integration (Chapters 1a, 1b and 2). Next, three main research lines are developed (Chapters 3, 4 and 5), followed by an extensive discussion on the achievements and challenges encountered throughout the research, as well as on ways forward and future exciting opportunities (Chapter 6). Chapter 7 provides a general conclusion.

Chapter 3 focused on exploring sediment-acoustic relationships in view of the available ground-truth data and proceeded with the predictive mapping of benthic substrates, based on a large multisource dataset for which a strategic surveying strategy allowed the production of a seamless, continuous backscatter map. Mapping of hard/heterogeneous substrate areas are of major importance in view of benthic habitat mapping, the identification of biodiversity hotspots and the designation and follow-up of macrobenthic communities and their status. The backscatter and bathymetry dataset were used to test the predictive performance of two statistical classifiers for the production of spatially-explicit models of benthic

substrate distribution: an unsupervised and a supervised classification routine using k-means partitive clustering and Random Forest Machine Learning classification, respectively. The discriminative ability of multibeam backscatter at 300 kHz was tested against two seafloor classification schemes, as well as various approaches to find the number of clusters (empirically and/or statistically). A range of accuracy metrics confirmed the strength and weaknesses of the automated classification predictions, based on backscatter alone, primary multibeam data (backscatter and bathymetry alone), and including a set of relevant morphometric and textural derivatives. Approaches were compared and evaluated, confirming the effectiveness of the supervised multivariate Machine Learning approach. The results of these analyses, and the strength of the predictions, were assessed by implementing a thorough protocol of error estimation based on the confusion matrix. Research comparing available classifiers are needed to confirm methodologies tested farther afield and explore the applicability of classification routines in different environmental settings with characteristic underlying data structures. Particularly, when prescribing classification schemes, awareness is needed on choosing appropriate aggregation level of a seafloor classification scheme since it may imply a loss of important environmental information.

Chapter 4 focused on investigating sources of variance caused by short-term, half-diel, tidal variability influencing multibeam backscatter measurements in the operational environment. This may have repercussions on how patterns and trends in serial backscatter datasets are interpreted, hence also on the assessment of naturally- and/or anthropogenically-induced variability of seafloor substrates. Particular attention was placed on the hydrological status and the acoustic transmission losses throughout the water column, having important implications on the retrieval of backscatter strength estimates that relate exclusively to the target of interest: the water-sediment interface. Ideally, results of echosounders across the globe should be comparable, enabling merging to produce large geographical coverage. In this regard, rigorous standards and control of the environmental drift on the backscatter measurements will be important to efficiently discern between changes that are due to seafloor substrate changes, from changes that are due to other processes and properties (e.g. water-column status, seafloor target-geometry, navigation heading). Three experiments were designed, targeting acquisition of multi-pass surveys over the duration of tidal cycles (short-term time scales) and covering three distinct seafloor types, representative of the main sediment classes in the Belgian Part of the North Sea. Building up a baseline that quantifies the sources of variance, their type (i.e. whether intrinsic or unwanted) and magnitudes, is important to understand the sensitivity of the hydroacoustic measurements as much as building libraries of backscatter signatures representative of given seafloor types.

Chapter 5 focuses on the innovative topic of acoustic change detection by testing and evaluating monitoring approaches that include relative calibration against a stable serial MBES backscatter dataset of a well-known location (effectively a

repeatability control). The dataset allows exploring pre-, post- and ensemble classification methodologies that are needed to gauge impact assessment. Controlling the measurements repeatability on natural reference areas is a pragmatic solution in mapping and monitoring programmes, in Belgian waters and beyond. In view of MSFD monitoring, approaches from the terrestrial remote sensing community are proposed: i.e. detection of spatio-temporal trends and quantification of the dominant signals of seafloor changes, such as persistence, gain, loss and from-to transitions between classified scenes.

Finally, Chapter 6 presents the overall Discussion, reiterating and expanding on the achievements and limitations identified in the research process, presenting research opportunities, and discussing the results in the context of acoustic seafloor classification and change detection, applicable to environmental monitoring. This chapter reiterates the importance of backscatter measurements' stability and repeatability and how its control allowed merging and comparing data in space and time, respectively. A critical evaluation of the data integration (classification) is proposed, identifying the salient issues, i.e. difficulties associated with finding the optimal number of classes, fitting of seafloor classification schemes, and fundamental physical properties challenging the discrimination potential of 300 kHz backscatter. Further argumentation on the ecological value and quantification of spatial uncertainty of the most accurate predictive models is presented as well. Finally, further insights are provided on the challenges faced by the seafloor mapping community when applying principles of acoustic change detection.

Chapter 1a

The Hunt-Lenox Globe ca. 1510.

Chapter 1a - Introduction

Seafloor mapping: sounding the unknown

1. Introduction

1.1 Seafloor mapping: sounding the unknown

Hic svnt leones et dracones (from Latin: here be lions and dragons) describes the mysterious cartographic gaps, or *mare in cognito*, reflecting our historical (Roman to medieval) perception of the ocean: the largest feature of planet Earth, covering over 70 % of its surface and accounting for an overall volume of approximately 1.3 billion km^3 (Charette and Smith, 2010). Beneath the Ocean surface, lies amongst the most cryptic, enchanting and unknown precincts of our planet: the seafloor (or benthic zone, from Greek *benthos*; "the depths", and coined by Ernst Haeckel in the late 19th century). As mankind mastered the capacity to travel on the water surface, depth measurements became of paramount importance to avoid running aground. The lead-line (a weighted and marked hemp rope – i.e. "a line and sinker system" – Fig. 1.1C) has been the first instrument employed by pioneering seafarers to remotely sense the depth of the water body and check under keel clearance. Early records of this technique can be dated as early as ~ 2000 BC with historical records of Egyptian "hydrographers" (Fig. 1.1A) employing such a tool (Bass, 1972).

For the next ~ 4000 years, followed by various modifications and improvements, the line length measured upon the arrival of the sinker to the bottom, remained the sole viable method to obtain depth measurements (Fig. 1.1B). Alternatively, the sinker was greased, allowing sticking of bottom sediments and giving an indication of the seafloor nature. Considerable improvements of the lead-line depth measurements were achieved by the piano-wire sounding system (Fig. 1.1D-E), invented by physicist and mathematician Lord Kelvin during the 1870's. This system consisted of a metallic piano wire drum, combined with a heavier sinker and a pressure gauge, which led to faster deployments and retrieval (even allowing "on the flight" deployments) and significantly reduced positional errors (about 20 m for a depth of 5 km compared to over 100 m using the former hemp ropes). Thanks to such improvements of these rudimental measurements, systematic charting of shallow maritime areas could be conducted for navigation safety purposes; moreover, some of the major bathymetric features (underwater geomorphology) were delineated. This led to the first hydrographic chart of the Mid-Atlantic Ridge, the first indication of the Mariana trench (at the time of the HMS Challenger recorded at a depth of ~8 km), identification of features in the Pacific Ocean such as the Challenger Tiefe and Deep and the Tuscarora Deep (Dierssen and Theberge, 2014).

Besides Leonardo da Vinci's pioneering experiments in underwater listening experiments of the 15th century (Gille, 1966), it was not until the early 1800's that scientists began to fully appreciate the utility and understand the characteristics of sound travelling in water. In 1826, scientists Charles Sturm and Daniel Colladon set out on Lake Geneva, Switzerland, embarked on two small vessels and navigated approximately 17 km apart from each other; one boat was equipped with an underwater bell and a flashlight, the other one with an underwater listening horn and

a chronometer (Lasky, 1977). The time elapsed between the flash and the bell's sound detected through the horn was measured to estimate a value of the underwater speed of sound; from this starting point the objective transmission characteristics of sound waves in water began to be fully appreciated by the scientific community. Shortly later (during 1880's), physicists Jacques and Pierre Curie discovered quartz piezoelectricity, the phenomenon by which pressure results in an electric potential, while carbon-button microphones were being invented by Emile Berliner (Klapholz, 1988), later contributing to the development of electroacoustic technology.

Based on this technological progress, the application of sound as an aid to navigation made its way into the early 1900's, when the Submarine Signal company (Howarth, 2015) began to equip buoys, lightships and vessels with sound emitting devices and hydrophones, drastically facilitating the entrance to harbours during periods of poor visibility. With the *Titanic* disaster (taking place in 1912 – Fig. 1.1F), the detection of obstacles hazardous to navigation such as icebergs, became even more critical. At this time, Canadian inventor Reginal Aubrey Fessenden pioneered the first electroacoustic transducer (named a Fessenden oscillator, and based on the same electrodynamic principle as aerial loudspeakers) and created a device that could both transmit and receive acoustic signals; hence able to detect the distance from the target and prevent a collision (Frost, 2001). The decisive impetus for technological development occurred shortly after, when during the First World War (and later during the Second one), the detection of German submarines (Fig. 1.1G) became a task of primary importance in the realm of maritime warfare. French physicist Paul Langevin and Russian engineer Constantin Shilovsky demonstrated that it was possible to transmit an acoustic signal (at 38 kHz) that could identify the presence of submarines, including the retrieval of their angular position and distance from the point of emission (Manbachi and Cobbold, 2011). This historical moment demarcated the advent of the first prototype of an active *sonar* system (Sound Navigation and Ranging).

By 1922, lead-line and piano-wire systems became obsolete overnight and echosounding became an increasingly accepted application. At the onset of this technological development, echosounding (taking from the Old French *sonder*, i.e. to measure) was a drastic improvement over traditional depth-measuring techniques (Fig. 1.1H), but it remained somewhat far from ideal as these early systems were designed to project very broad beams (in the range 30-60°). As a result, a considerably large portion of the seafloor could be ensonified, implying that the first returning echo could originate anywhere within an area approximately half to one time the water depth. The lack of angular resolution of these broad beams resulted in a fairly inaccurate depiction of seafloor topography and the target detection underneath the vessel (i.e. at nadir) could be somewhat biased due to echoes potentially coming from adjacent slopes. Between the two World Wars, particularly at the beginning of World War II, the maritime warfare and engineering sectors worked towards the development of narrow-beam echosounders and by this moment, vessels and submersibles were starting to be equipped with outward and/or downward looking target/depth finding acoustic systems. With narrow-beam echosounders, the accuracy of bottom detection became significantly more reliable. However, the area ensonified through narrow

beams is intrinsically very small, resulting in a spatially sparse sampling of the seafloor. This limitation had tremendously progressed during the 40's, when echosounder components were starting to be installed/mounted in such a way so as to transmit and receive multiple narrow beams and insonify a large portion of the seafloor, composed of several tightly spaced ensonified areas: together referred to as a swath (Fish and Carr, 1990 - Fig. 1.1H).

Owing to a significant and rapid development of electroacoustic transducers, electronics, and finally computers and digital processing, the late 1970's witnessed the establishment of today's generation of active sonars which became available for civilian and scientific applications: multibeam echosounding systems (MBES). De Moustier (1986) introduced the application of this technology to remotely characterise the seafloor nature (i.e. not only its depth) from the measurement of the backscatter level over the angular range made possible by MBES technology.

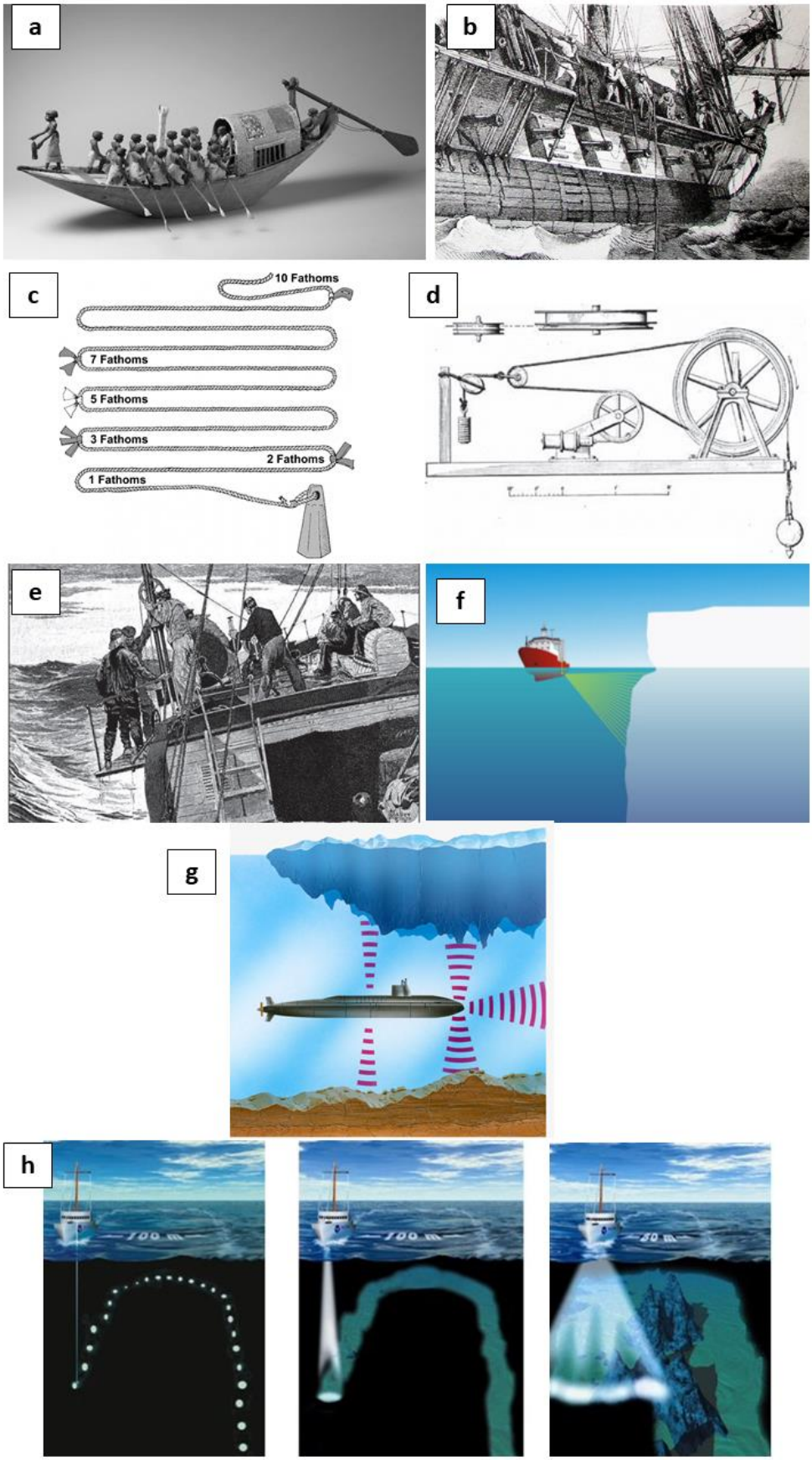


Figure 1.1 – A brief history of seafaring, depth-measuring instrumentation. Caption continues on next page.

A) Model of an Egyptian vessel retrieved from Meket-re's tomb, buried in Thebes in 2000 BS. From the Metropolitan Museum Arts, New York, USA, collection (from: https://images.metmuseum.org/CRDImages/eg/original/DP249000.jpg). B) In the 1840's British naval expeditions employ the lead-line (in C) to map e.g. the Gulf of Mexico (from: http://larrymayer.net/history-of-bathymetry-early-methods/). C) The lead-line displaying various intervals (knots and textile material) used to measure depth in fathoms (1 fathom = 1.82 m – from: https://seahistory.org/wp-content/uploads/Lead-line-Fathoms.jpg). D) in the 1870's, the piano-wire sounding system is designed. A sketch of the original design is shown (from: Dierssen and Theberge, 2014). E) An original image from "The United States Fish Commission" by Richard Rathbun; Century Magazine 1892. "Sounding the abyss with piano-wire "showing a modification of the former piano-wire sounding machine (now Sigsbee sounding machine – from: Dierssen and Theberge, 2014). F) Representation of ice-berg detection by sideward looking sonar (from: Kimball and Rock, 2011). G) A schematic representation of naval active sonar detection of submarines and obstacles (from: https://img1.cgtrader.com/items/706508/d44bc27972/large/german-u-boat-type-vii-3d-model-max-fbx.jpg). H) Evolution of depth-measuring instruments from lead-line, to single beam and multi beam echosounding systems displaying the kind of bottom coverage obtained by the three methods (from : https://noaacoastsurvey.files.wordpress.com/2015/07/surveying.jpg).

The next section provides a general background to the modern use of acoustic in seafloor mapping, introducing this nascent discipline, the multibeam echosounding technology and associated data types, and finally the topical issues which motivated this PhD research. Following, the socio-legal framework and objectives of the broader project within which this research is conceived is described, and the research questions are stated. A more detailed introduction to principles of multibeam echosounding and seafloor-acoustic interaction are in turn presented in Chapter 2.

1.2 State of the art in seafloor mapping: research background

Besides aiding navigation, detecting fish or submarines and providing the optimal means to measure the great depths of our planet, echosounding has become a keystone technique for scientists studying the seafloor in a rapidly changing and anthropogenically threatened marine ecosystem. The Anthropocene era sees the human dependence on resources and ecosystem services provided by the marine environment as vertiginously growing with the human population exceeding 7.5 billion individuals (UNFPA, 2018). A global synergy of pervasive anthropogenic pressures threatens marine ecosystems through large-scale coastal and offshore infrastructural developments, intense routed navigation, invasive and intensive commercial fishing practices and widespread mineral resource extraction (Halpern et al., 2008). A general global realisation of these overarching impacts resulted in numerous efforts being initiated towards the exploration, understanding and mitigation of human impacts and the implementation of a more responsible stewardship of the marine environment (i.e. the initiation of nationally funded seafloor mapping programmes – see Chapter 3). This has led to a general shift from single-resource (i.e. single species) management to Ecosystem Based Management (EBM) and Maritime Spatial Planning (MSP) (i.e. large geographical areas and “place-based” management), recognising the full extent of complex interactions of marine organisms with their abiotic environment (Curtin and Prellezo, 2010). A first and important step towards a better management (monitoring) of marine natural resources is the acknowledgment of their spatial extent and organisation (Diaz et al., 2004). The spatial nature of the required information largely relies on the application of underwater acoustics (aka hydroacoustics) to seafloor mapping as the fundamental analytical application underpinning the interpretation of this largely unknown and often operationally challenging environment, possibly at multiple spatio-temporal scales.

Recent developments in hydroacoustic remote sensing instrumentation, in particular the specific design of multibeam echosounders (MBES), had major implications in the field of seafloor mapping since the late 1980’s (De Moustier, 1986). Fundamentally, the advantage of using MBES over former echosounding technologies is in the ability to co-register accurate (i.e. motion-and refraction-compensated), continuous (i.e. high-coverage) and detailed (i.e. high-resolution and precisely georeferenced) bathymetry and backscatter data (Fig. 1.1H). Bathymetry is the underwater equivalent of terrestrial topography, used to obtain an understanding of the morphology and three-dimensional organisation of the seafloor continuum (Lecours et al., 2016a). The backscatter strength (or more generally referred to as *reflectivity* and now referring to the relative intensity of the returned signal as a complement to its flight-time used for bathymetry) was only considered as a sonar by-product until recently, despite it being a direct proxy of the nature of the water-sediment interface. This physical phenomenon is the principle exploited in the design and operation of all underwater echosounding instruments: the emitted sound wave will return to the echosounder following a complex interaction with the seafloor which is inherently “able” to reflect a measurable part of the acoustic energy, making sounding

practically feasible. How much of this energy will return to the sonar depends on the seafloor type i.e. on its characteristic impedance contrast, interface roughness and volume inhomogeneity (see *Backscattering from the* seafloor in Chapter 2), but also on the signal incidence angle on the seafloor interface and on the acoustical frequency (Lurton, 2010). As a result, the notion that this quantity could be derived from the same depth-measuring echo, and used to remotely characterise the benthic substrate, quickly made its way in the interested scientific community (e.g. Hamilton, 1980). The quality of depth data is obviously central to hydrographers, leading to developing well-established standards of acquisition and processing and uncertainty budgets (i.e. the International Hydrographic Organisation - IHO, 2008). Contrariwise, backscatter data have remained largely under-exploited due to the inherent technical constraints, mostly relating to limited digital signal processing and the design and development of electro-acoustic components which came to technological maturity only in the 1990's (Kenny, 2003; Lurton and Lamarche et al., 2015; Malik et al., 2018). Today, acoustical backscatter is recognised as a fundamental and unique kind of datum for the characterisation of the seafloor nature, receiving increasing attention by several disciplines in the marine sciences which aims at reaching standards in calibration, acquisitional and processing as in the terrestrial remote-sensing realm (Buck, 2000; Eleftherakis et al., 2018; Figure 1.2) where a considerably less challenging operational environment constraints the success of *in situ* measurements.

Both pioneering (e.g. Hamilton, 1980 and following work) and more recent investigations (e.g. Goff et al., 2000; Collier and Brown, 2005; Ferrini and Flood, 2006; De Falco et al., 2010; Gaida et al., 2018) showed the strong theoretical and empirical relationships between backscattered echo characteristics and measurable sediment characteristics. Establishing, when possible, direct relationships (for example with the sediment grain size) improving the general appreciation that backscatter data can yield qualitative and quantitative information of the seafloor nature, relating to its texture and composition, and subsequently, potential for seafloor type classification (homologous to Land Cover Land Use LULC applications by the terrestrial remote sensing community). Conceptually, a "chain of proxies" towards the production of thematic maps has henceforth been established: (1) backscatter data relates to the sediment nature (Lurton, 2010), (2) the benthic substrate is recognised as a fundamental physical support of the benthic habitat (McArthur et al., 2010), (3) hence interpretation of backscatter data can indirectly inform the distribution of biodiversity (Diaz et al., 2004). Lastly, (4) benthic organisms are indicative of the ecological status of an area (whether in favourable conditions or not - Muxika et al., 2005). Of course, backscatter data cannot be the sole means of biological inference (with the exception of biogenic/structuring species) and requires integration with an array of environmental datasets (such as modelled spatial data and ground-truth information), which, to certain degrees of confidence and generalisation, can lead to the production of a variety of purpose-made seafloor maps, including benthic habitats (e.g. Brown et al., 2012; Brown and Blondel, 2009; Che Hasan et al., 2014; Lacharité et al., 2018).

With regard to benthic habitats, multibeam technology has drastically improved their detection and monitoring (e.g. van Rein et al., 2011; Rattray et al., 2013). While in terrestrial landscapes and tropical maritime environments habitats can be studied and identified based on salient biotic characteristics of which importantly foundation and structuring species such as vegetation (i.e. mappable by remote sensing approaches e.g. Macroalgae - Kruss et al., 2008, 2011, 2015, 2017), for the most part, it is the identification of non-living features (topography, patches, sediment types such as shell hash, gravel etc.) that dictates our ability to identify links with the benthic life (Zajac, 1999; Pittman et al., 2011), especially in predominantly sedimentary marginal Northern Atlantic continental shelf environments.

The need to inform policy-making with scientific and accurate spatial data (spatio-temporally describing the submerged environment in an explicit manner) and the rapidly growing volumes of MBES data being collected, have led to the emerging discipline of Acoustic Seafloor Classification (ASC - Anderson et al., 2008) that developed an array of strategies dedicated to the integration of hydroacoustic and ground-truth observations (Brown et al., 2012). This discipline seeks to establish univocal links between remotely sensed data and the real world, producing detailed cartographic and thematic models (i.e. maps reducing spatial information, often of a complex multivariate structure, into more easily interpretable categorical themes) containing keystone geological, sedimentological and biological information (e.g. Fig. 1.3) and validated by *in situ* sampling approaches (i.e. ground truthing). The innovative character of this nascent and rapidly developing discipline is in the shift from the application of interpretative (expert-judgment-driven) to objective methods (i.e. statistically-driven) in the classification and interpretation of the acoustic data.

As noted by Butman et al. (1992) in this respect: *"Remote sensing mapping techniques are essential to adequately determine the complex spatial variability of the bottom morphology and sediment texture: accurate maps cannot be prepared from the analysis of sediment grab samples alone."*

As later described in more detail (Chapter 2), the developing data-integration methodologies in the acoustic seafloor classification realm can be summarised by three central factors: (1) the data type, whether processed as hydroacoustic signals (geoaocustically and/or empirically modelled angular responses) or images (data gridded into sonographs/mosaics and terrain models); (2) the target of classification, whether categorical (thematic / semi-quantitative / discrete) or continuous (numerical / quantitative / continuous); and (3) the inclusion of ground-truth validation data in the integration process allowing applying a supervised classification (with *a priori* information) instead of an unsupervised classification (no *a priori* information of the underlying seafloor composition).

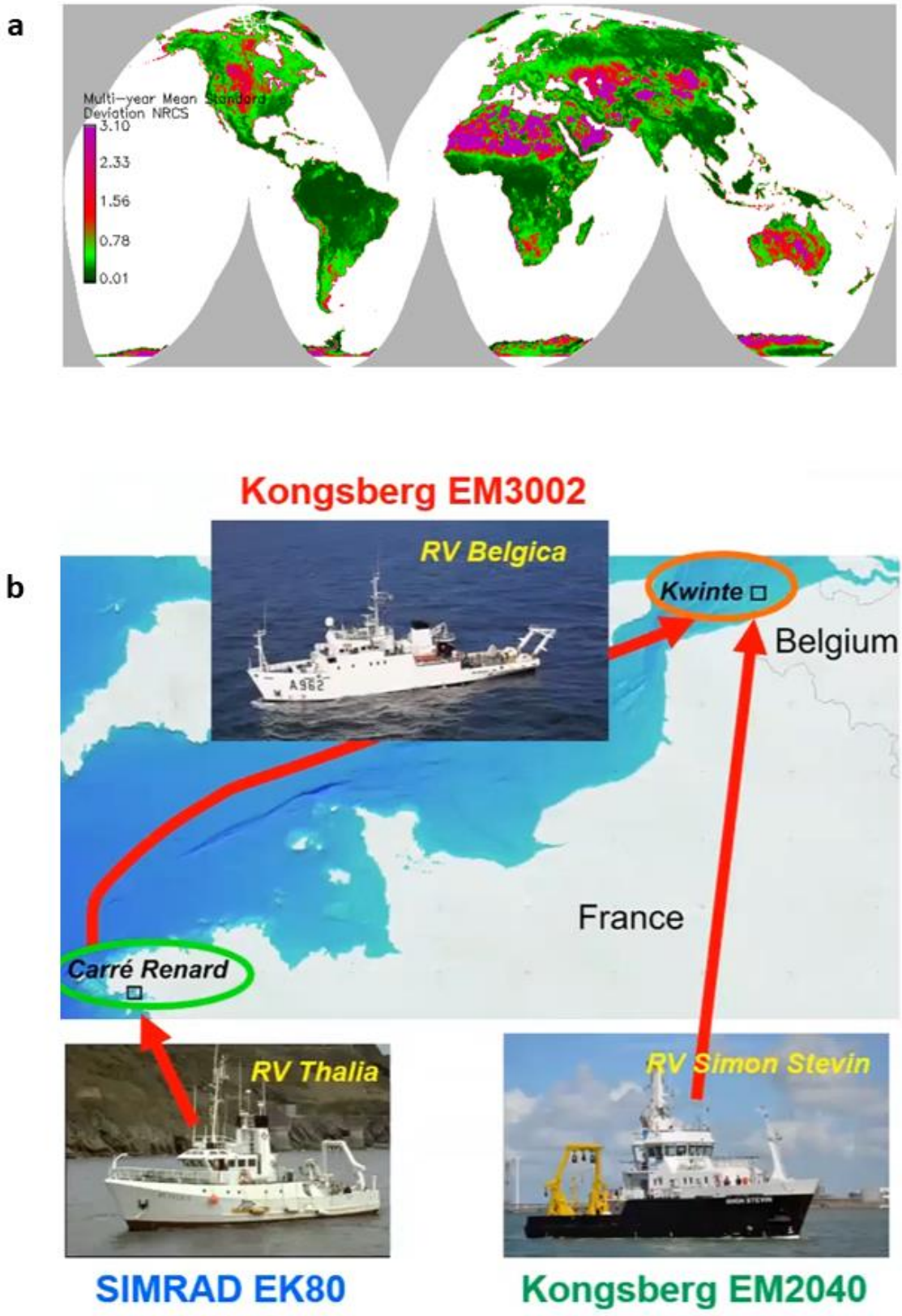


Figure 1.2 – Remote sensing sensors calibration from space and from sea surface. Caption continues on next page. (a) From space using a Synthetic Aperture radar (SAR) scatterometer and (b) in the ocean; propagating the calibrated measurements of one sensor to those of non-calibrated ones based on a common survey. In (a) the forest canopy of the Amazonian rain forest, maintains a year-round

stable average reflectivity value of -10 dB and has a standard deviation of 0.01 dB, making the optimal calibration target onto which the terrestrial remote-sensing community relies for several decades (from Buck, 2000). In (b) the situation at sea is bounded to different logistics: MBES sensors have a considerably smaller field of view compared to a sensor in orbit and the calibrated measurements from one vessel and echosounder have to be propagated to those of other platforms. This requires the discovery of seafloor natural and stable reference areas and "shared" oceanographic missions between bordering countries, for example at the European level. As exemplified in (b), the calibrated measurements of a French vessel (*RV Thalia* -IFREMER) are propagated to two neighbouring Belgian *RV*s based on common measurements on stable reference areas (b – from University New Hampshire web seminar by Professor Xavier Lurton- https://vimeo.com/299732070). It must be noted that the cross-calibration propagation can occur on any area (with favourable conditions) as long as one of the echosounders has undergone an absolute calibration from which biases can be estimated. The well-known and stable natural reference seafloor area can represent a source of control and stability check for a system in use and for relative calibration propagation between different systems (again by survey overlap).

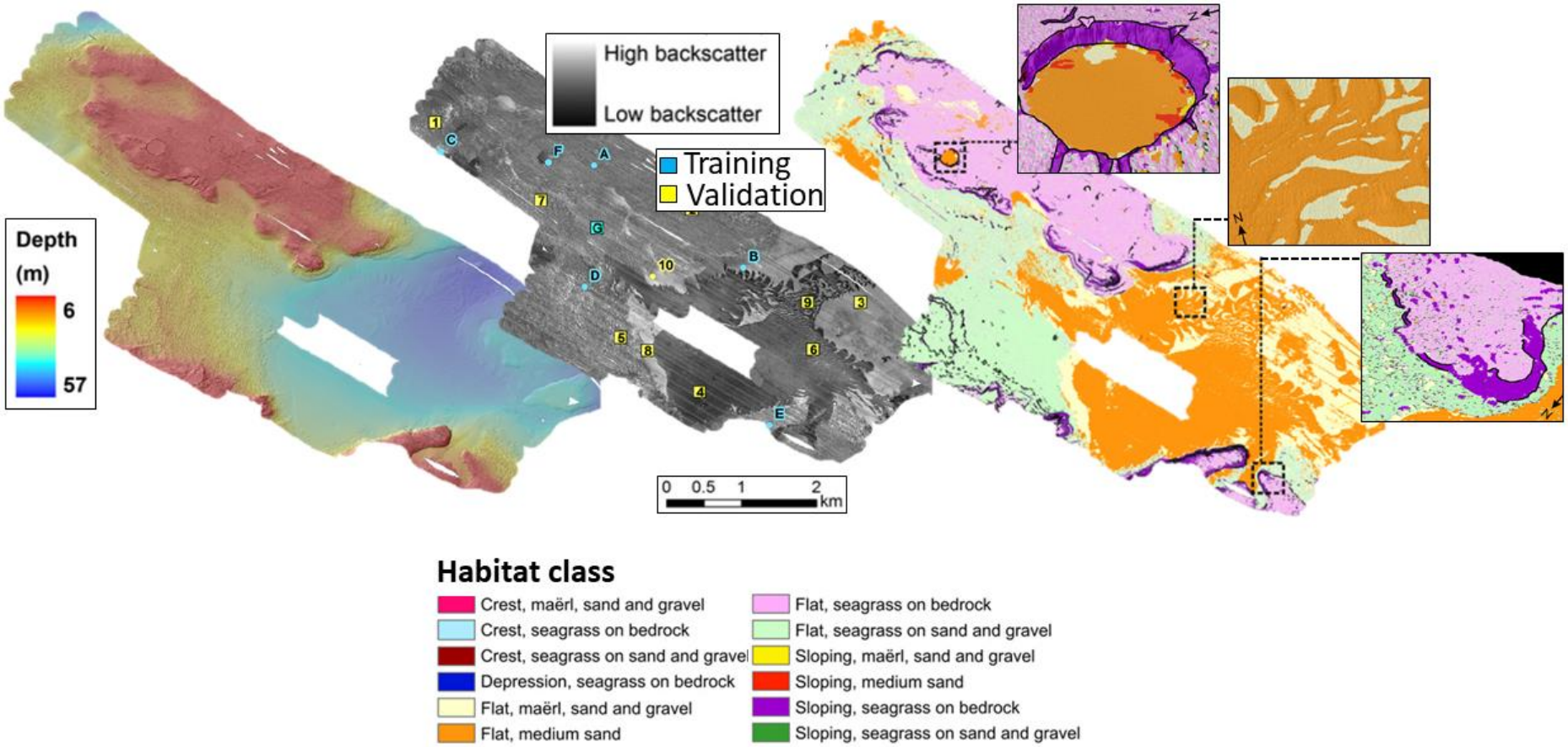


Figure 1.3 - Example of the integration of MBES hydroacoustic datasets. From left to right: gridded bathymetry, backscatter, and spatially-explicit thematic model resulting from the integration of geophysical and ground-truth data. The legend provides the means to interpret the predicted seafloor classes of the coastal waters of the Maltese islands (modified from Micallef et al., 2012).

Interdisciplinary research efforts dedicated to the classification of MBES backscatter data into meaningful thematic products have developed at a fast pace over the past decade. Oppositely, there is a paucity of efforts quantifying spatio-temporal patterns of change in benthic substrates and habitats using backscatter measurements. However, the breakthrough this approach can bring into assessing environmental status of the seafloor justifies the considerable costs associated with repeating surveys and solving the technicalities involved in controlling the stability and repeatability of the measurements (see *Backscatter calibration, repeatability and standards in acquisition and processing* in Chapter 2).

A selection of noticeable investigations on acoustic seafloor classification in space (mapping) and time (monitoring) using multibeam backscatter data are presented hereafter.

1.2.1 Multibeam backscatter for benthic habitat mapping

Todd et al., (2000) and Kostylev et al., (2001) conducted seafloor mapping studies in the Browns Bank on the Scotian Shelf of Atlantic Canada, acquiring ~ 3000 km^2 of MBES bathymetry and backscatter data integrated with sedimentological and biological samples by means of a manual classification approach (i.e. the former generation, geologically-oriented manual digitisation of gridded backscatter datasets). Via ordination and clustering techniques, they could predict the distribution of six benthic habitat classes, including habitats of noticeable conservation importance (i.e. soft coral and sea cucumber community) as well as of commercial importance (i.e. scallop bivalve habitat), enormously enhancing the understanding of the bank ecology, sedimentology and geomorphology. They found backscatter measurements to be fundamental in the identification of harder and softer substrates, hence aiding the interpretation of the physical support of various benthic biota and sedimentary processes.

In the same waters of St. Anns Bank area, Lacharité et al. (2017) integrated a large multi-source, non-overlapping set of MBES backscatter data covering altogether 2870 km^2. Semi-automated image-analysis techniques were applied to backscatter and bathymetric data to predict the distribution of seven habitats ranging from predominantly bare mud with seapens to gravel with crustose coralline algae and crinoids. This study underlined important emerging challenges associated with merging of multi-source backscatter surveys in space and time. The need for dedicated surveying strategies and standardisation of operational procedures were advocated towards the achievement of backscatter measurements on a comparable scale; improving the assessment of temporal changes, the spatial comparison of similar seafloor settings in different survey areas and/or within single regions with varying surveys.

In the Mediterranean Sea, Micallef et al. (2012 - Fig. 1.3) and De Falco et al. (2010) present MBES backscatter facies relating to habitats of primary conservation

importance, such as seagrass and maërl cover: important benthic features promoting biodiversity.

Montereale Gavazzi et al. (2016) used very-high-resolution backscatter images to identify the signatures of single demosponges in a tidal channel of the Venice Lagoon, contributing to the discovery of a new habitat for this increasingly more explored environment (Madricardo et al., 2017; 2019). Besides habitat mapping (requiring integration of the hydroacoustic data with biological samples), MBES backscatter has been the centre of attention in studies concerning the prediction of benthic substrates, for which strong empirical links between both data types have been established by many.

1.2.2 Multibeam backscatter for prediction of benthic substrates

In the North Sea, north-eastern coast of England, Stephens and Diesing (2014) explored and tested the predictive accuracy of an array of automated image-classification approaches to predict four substrate classes ranging from mud to coarse sands and rock within a ~ 5 km^2 study area. Their study concluded that backscatter data were *"by far the most important"* variable in improving model accuracy and discriminating between substrate types compared to a range of spatial layers derived from the primary MBES data (bathymetry and backscatter).

Lamarche et al. (2011) presented a regional and calibrated backscatter dataset acquired over ~8500 km^2 in the Cook Strait area between the North and South Islands of New Zealand. The study demonstrated the potential of both image- and signal-processing techniques to unsupervisedly (i.e. without recurring to direct sampling) capture the broad geological and sedimentological nature of their study area.

Moving to research on mineral resources of commercial interest, Gazis et al. (2018) integrated backscatter and bathymetric derivatives using an innovative machine learning approach predicting the continuous distribution of manganese nodules within an extraction concession zone in the Pacific Ocean, and Naudts et al. (2008) related patterns in backscatter imagery to seepage of methane gas in an area of the Black Sea.

1.2.3 Multibeam backscatter for monitoring changes in benthic substrates

Besides establishing links between MBES backscatter data and benthic substrates and features, repeated surveys pose the remarkable opportunity to detect patterns of change in the distribution of the identified classes, raising the potential for an array of environmental monitoring applications (Fig. 1.4).

An important study by Urgeles et al. (2002), demonstrated the potential of broad (> 1 km) scale monitoring of benthic substrates in the Saguenay Fjord (Quebec, Canada) based on multibeam backscatter measurements (Surveys where acquired before and after a massive flood deposit event in 1996). Using a natural reference area to inter-calibrate the backscatter surveys, they present a three-year change detection study identifying the large scale geological and sedimentological patterns of change in

the fjord. By collecting ground truth data complementary to each survey, the study could relate part of the patterns of change identified in the backscatter imagery to patterns of bioturbation, specifically to the recovery of biogenic structures following severe smothering caused by sediments brought by the flood event. Urgeles et al. (2002) stepping stone investigation demonstrates how significant seafloor ecosystem changes can be monitored from multibeam backscatter measurements.

A noticeable study is that of Rattray et al. (2013) wherein a bi-temporal change detection analysis (1-year interval), based on automated classification of backscatter and bathymetric derivatives, could identify patterns of change in ecologically noteworthy habitats such as between Kelp-dominated areas and barren sediment grounds within an area of 18 km^2 in Western Australian waters (Fig. 1.4A). Their study concluded that MBES-based surveying was the optimal solution to map and detect changes in seafloor habitats, particularly beyond the penetration depths of optical remote sensing instruments. Their study observed considerable variation in backscatter values between the uncalibrated and unstable serial surveys, making the direct comparison of backscatter measurements unfeasible. Assuming appropriate radiometric and geometric corrections have been implemented (see later paragraph *Processing and correcting MBES* backscatter) they argued in favour of post-classification approaches to mitigate the important issue of direct data comparability. They equally stressed the need for calibration and/or stability control of the measurements to effectively and directly exploit the backscatter measurements in the context of change detection.

The utility of seafloor substrate monitoring using stable and controlled backscatter measurements is exemplified in the works of Roche et al. (2017) and co-workers from the Continental Shelf Service of Belgium (CSS). In fulfilment of the implementation of legally mandated monitoring of geomorphological and seabed changes in areas of marine aggregate extraction (namely sand and gravel) in the Belgian Part of the North Sea (BPNS), they have acquired a unique corpus of serial backscatter datasets since 1999. Their studies demonstrate that trends in the average backscatter response from serial datasets within selected areas can be related to trends in sedimentary changes, in this case resulting from the activity of trailing suction hopper dredging vessels (Fig. 1.4B).

Recent groundwork in the Bay of Biscay and the Celtic Sea by Fezzani and Berger (2018) demonstrated an approach to detect changes of the water-sediment interface using calibrated backscatter data acquired within a sandy seafloor area between 2010 and 2015 and processed as acoustic signals (i.e. angular responses describing the variation of backscatter intensity as a function of the angle of ensonification on the seafloor – see Chapter 2 for details and 3 and 4 for applications – Fig. 1.4C). Here, changes in calibrated (i.e. directly comparable) backscatter data could be related to changes in the interface target geometry, namely the geometric reorganisation in the orientation of sand ripples. Besides the potential of angular

response backscatter in detecting changes of the water-sediment interface, this study presents a unique corpus of calibrated backscatter data acquired by a single sensor at a regional scale (the entire Bay of Biscay) and used for seafloor-type and habitat mapping, with a particular focus on demersal fish-habitat: useful to improve management of commercial fisheries for this area. The study suffered from a paucity of ground truth data to validate the results in terms of substrate/habitat type, pointing at the uttermost importance of acquiring concurrent samples. However, the modelling applied to the angular response backscatter allowed identifying the salient sediment characteristics, providing an unsupervised application for a general but physical and geologically meaningful interpretation (see Acoustic signals and images, Chapter 2).

Overall, the multi- and trans-disciplinary nature of these investigations are a clear indication of the promising utilisation of MBES backscatter data in the interest of mapping and monitoring sedimentological, geological, hydrological, geochemical, biological and ecological processes.

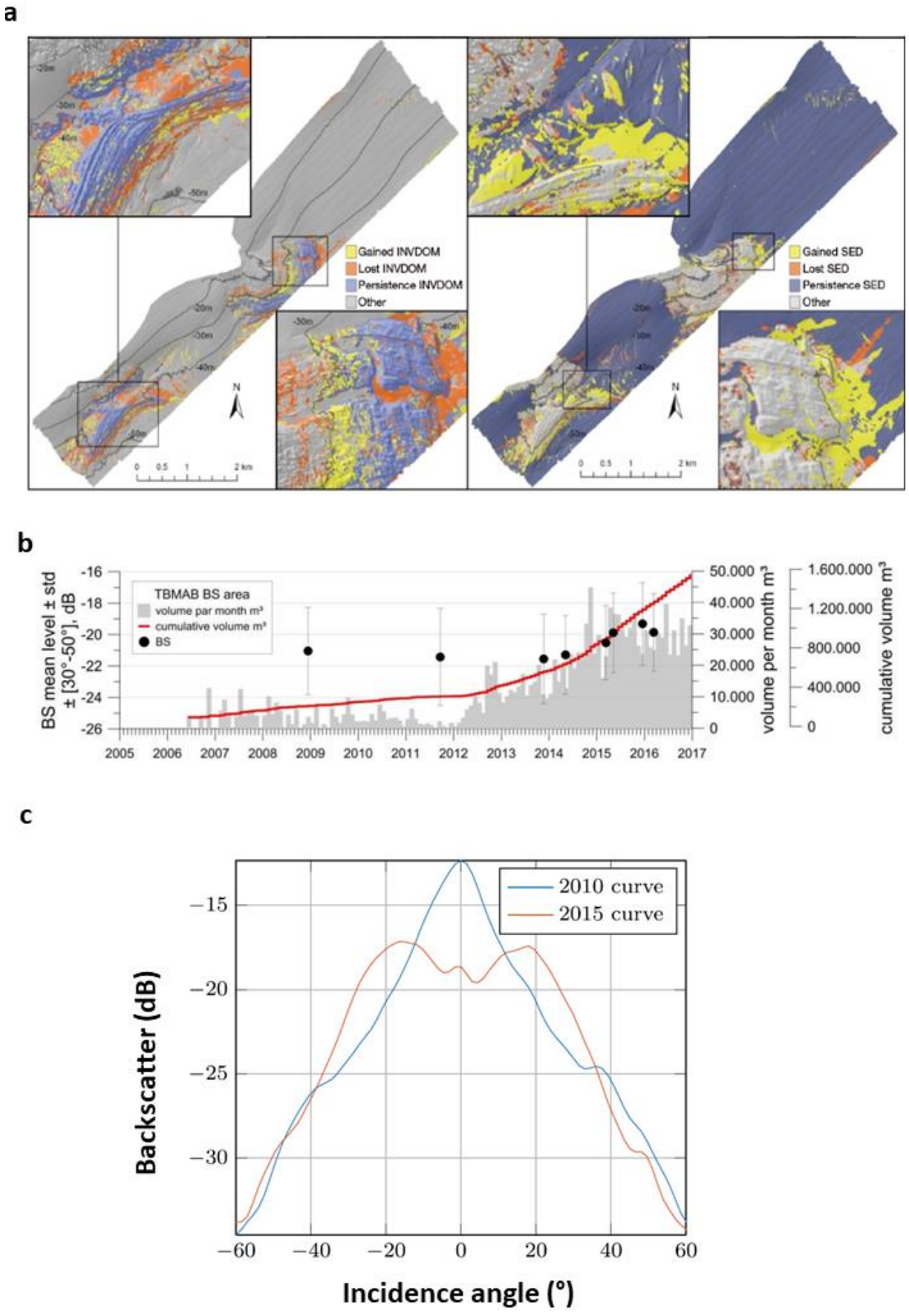


Figure 1.4 – Example of serial MBES datasets dedicated to environmental monitoring. Caption continues on next page.

a) Example of a bi-temporal post-classification change detection showing patterns of persistence, gain and loss of two habitat classes. Note how a loss of the INVDOM class (left quadrant) results in a gain of the SED class (right) (modified from Rattray et al., 2013). b) Example of a time series of stable and repeatable backscatter measurements directly used to observe the sedimentary evolution of a sand extraction monitoring area in the BPNS. The backscatter time series (average values within the region of interest – black points) is presented in complement to the cumulative extracted volume (red line) showing a correspondence between the two data series. Here, the modifications of the substrate nature are related to the direct removal of the extracted sand, the establishment of a pit and the accumulation of bioclastic detritus (shell thanatocoenosis) increasing the overall reflectivity of the study site (from Roche et al., 2017). c) A change in values and geometry of calibrated backscatter angular responses is indicative of local seafloor changes (modified from Fezzani and Berger, 2018).

In this context, this doctoral thesis endeavoured in addressing research challenges associated with the utilisation of multibeam backscatter measurements in the framework of mapping and monitoring the submerged environment and aimed at advancing the setting-up of a seafloor-mapping strategy for assessing environmental changes in the Belgian Part of the North Sea. In the next sub-chapter (Chapter 1b), the specific background to the socio-legal and wider project's framework within which this PhD research is rooted is presented, followed by the document's structure and the specific research questions herein dealt with.

Chapter 1b

1.3 Thesis framework and research questions

1.3.1 Thesis sociolegal background

This doctoral research falls under the framework of the INDI67 project, funded by the Belgian Scientific Policy Office (Contract Grant Nr. BR/143/A2/INDI67) that aimed at developing tools supporting the monitoring objectives of the European Marine Strategy Framework Directive (MSFD Directive 2008/56/EC) descriptors D6 and D7 on "Seafloor Integrity" and "Hydrographic Conditions", respectively. Five working packages were identified of which one is dedicated to "Advanced mapping of seafloor/habitat types"; strictly related to D6 and to which this research is bounded. Recognising in full the spatial nature of maritime anthropogenic activities and that knowledge of the seafloor composition and of its spatio-temporal evolution is of great relevance to monitor possible human impacts on benthic habitats, the MSFD legislation mandates European Member States to cyclically (on a six-year basis) map and monitor their marine waters towards the achievement of a "Good Environmental Status (GES)" (Borja et al., 2013, for a review). Within this framework, twelve GES descriptors were put forward for which each EU Member State defined indicators with associated monitoring programmes.

In D6 terms, GES is achieved when the seafloor integrity is such that *"the structure and functions of the ecosystems are safeguarded and that the benthic ecosystems are not adversely affected"* (see Rice et al., 2012 for a review). Physical loss and physical disturbance are primary criteria for the assessment of seabed integrity and mapping of human pressures is needed to support the biological evaluation of the status of marine benthic habitats (Commission Decision 2017/848, 17/5/2017). Physical loss is defined as a permanent change of the seabed for two MSFD cycles (12 years) or more, and physical disturbance as a change in the seabed that can be restored if the activity responsible for such physical pressures ceases to exist. ICES (2019) advised on extending this definition stating that physical loss is any human-induced permanent alteration of the physical habitat from which recovery is impossible without further intervention. The effective indication of 'loss' or 'disturbance' per activity is open to interpretation and may vary from one Member State to another (see Kint et al., 2018 for a discussion).

For Belgian waters, physical loss and disturbance were mapped at different scales (Van Lancker et al., 2018): (1) Spatial and temporal mapping of all pressures resulting from human activities. (2) Mapping and monitoring of spatial management actions that aim to reduce pressures related to fishing. (3) Mapping and monitoring of the occurrence of large-scale benthic habitats on a BPNS scale using acoustic measurement methods and sampling. (4) Trend mapping of the evolution of morphology and the type of substrate in gravel beds in two test zones using acoustic measurement techniques, sampling and observations.

Points (3) and (4) form the background of this doctoral thesis and relate to the following environmental targets that were specified during the Initial Assessment phase (Belgian State, 2012).

1. The areal extent and meso to broad-scale (≥ 100 m; ≥ 1 km mapping unit) distribution of the European Nature Information System (EUNIS) level II Habitats (sandy mud to mud; muddy sand to sand and coarse sediments), as well as of the gravel beds, remain within the margin of uncertainty of the sediment distribution with reference to the Initial Assessment map (Belgian State 2012).

2. Specifically related to the ecologically noteworthy gravel beds, it is specified that the fine-scale (< 10 m mapping unit) ratio of the hard (gravel) substrate surface area to the soft (sand) substrate surface area must not show a negative trend.

Prior to implementing legal objectives and conducting the monitoring, setting up a seafloor mapping strategy is required, as well as resolving challenges inherent to acquiring, integrating and comparing datasets (generally introduced in Chapter 1a). Therefore, the project initially formulated a set of general objectives to advance mapping and monitoring of seafloor substrate/habitat-type listed hereafter:

A. Estimate the precision, sensitivity and repeatability of the backscatter measurements and measuring devices to detect spatial variability in seafloor substrate/habitat type, including the definition of best practices in surveying and ground-truthing the acoustic signal.

   Here, the research targets the investigation of the discrimination potential of multibeam backscatter in view of mapping the main granulometry classes of benthic substrates in the Belgian region: mud, sand and gravel.

B. Quantify external sources of variance responsible for unwanted fluctuations in the backscatter measurements.

   Here, the research targets the investigation of external sources of variance, setting up dedicated experiments quantifying magnitudes of variability (dB), with a focus on the hydrological status of the surveying environment.

C. Set up methodological frameworks towards the detection of patterns of changes in seafloor substrate type;

Here research is needed to test and define approaches to detect changes in seafloor composition, based on integrated hydroacoustic and ground truth measurements.

D. Innovate in collaborative seabed mapping.

The establishment of a Community of Practice (CoP) is here targeted involving the main seabed mapping parties in Belgium (i.e. the Continental Shelf Service of the Federal Public Service Economy, the Flemish Hydrographic Service, the Operational Directorate of Natural Environments of the Royal Belgian Institute of Natural Sciences, Belgian Navy and Flanders Marine Institute). This is needed towards the optimisation of surveying and monitoring efforts.

1.3.2 Thesis structure and research questions

Motivating this doctoral thesis is the recognition that the world-ocean floor faces increasing human pressures by a multitude of intensive and pervasive economic activities. To date, our understanding of this environment and of our impacts at local, regional and global scales is progressively improving but remains scarce. Global awareness of such a realisation resulted in a number of initiatives dedicated to discovering, evaluating and monitoring marine areas affected by such pressures. Multibeam echosounders have revolutionised our ability to visualise the submerged environment improving our attempt at a more responsible stewardship of marine resources. State-of-the-art backscatter measurements in particular, have drastically improved our ability to characterise the seafloor nature at continuous scales by the proxy approach, crucial in predicting biodiversity besides other numerous uses.

The work herein presented addresses various issues and applications of acoustic seafloor classification pertinent to an improved understanding of the distribution of benthic substrates and habitats and the advanced means to characterise them in space and time. To that end, this thesis is structured into a set of three independent, yet related investigations placed around the use of remote and direct (ground truth sampling) observations of the seafloor using some of the contemporary technology and exploring advanced techniques deriving seafloor sediment maps and detecting patterns of change. The studies test assumptions of data stability, repeatability and discriminative ability. Key aspects dealt with relate to identifying objective data-integration routines, assessing strength of association between the remotely sensed data and the measurable physical characteristics of the ground truth data and testing the sensitivity of the backscatter measurements to environmental variability, having implications towards mapping and monitoring applications.

The thesis is structured as follows:

- **Chapter 1a: Introduction**

Chapter 1a covered a brief historical background to the realm of underwater exploration, continued with a presentation of modern seafloor mapping and applications, introducing the general topical issues which motivated this research.

- **Chapter 1b: Thesis framework and research questions**

Herewith, **Chapter 1b** detailed this thesis background project's sociolegal framework and objectives. Hereafter, the research questions dealt with in this research are specified.

- **Chapter 2** (Multibeam echosounding: state of the art of hydroacoustic remote sensing), proposes an in-depth presentation of the multibeam echo sounding today's technology and an introduction to the principles of the seafloor backscattering phenomenon is provided. This includes the presentation of data integration routines, describing the general methodological approaches, of the data types' reduction, analysis and interpretation, and a literature review of the inherent research. A general introduction to the Belgian Part of the North Sea, this thesis study area, is given in Chapter 3, and in each chapter for the study areas therein investigated.

- **Chapter 3**, relating to the overall objective A, is on "Integrating multi-source multibeam and ground-truth data to seamlessly map continental shelf substrate types: Application to the Belgian Part of the North Sea" and was driven by the following research questions:

    - Which kind and how many sediment classes can be discerned (i.e. what is the discriminative ability) in the backscatter data of a sediment-dominated continental shelf area?
    - How can MBES hydroacoustic and ground truth data be effectively integrated to accurately predict sediment type and produce spatially explicit models of sediment distribution?
    - Are there trade-offs between classification scheme and predictive accuracy of the models?

To address these questions, MBES and ground-truth data, acquired over eight oceanographic campaigns in the BPNS, provide the background to setting up a seafloor mapping strategy and test data integration approaches. Unsupervised clustering and supervised machine-learning classifiers (k-means and Random Forest) were tested for their predictive accuracy, including the effect of two substrate classification schemes (how many classes can be mapped and how does the prescribed classification scheme fit the data?). The devised seafloor mapping strategy and the variety of sediment types and study areas surveyed provides the optimal case

study to test acoustic seafloor classification and test the discrimination potential of backscatter data. To constrain further the questions, empirical relationships between MBES backscatter and grain-size attributes were investigated. Finally, the modelling approaches allowed to objectively capture a number of surficial sediment types, producing state-of-the-art spatially-explicit substrate models.

- **Chapter 4**, relating to overall objective B, is entitled “Insights into the Short-Term Tidal Variability of Multibeam Backscatter from Field Experiments on Different Seafloor Types” and is a published paper driven by the following research questions:

    - While surveying, which external factors are responsible for unwanted fluctuations of the backscatter measurements?
    - How can these be experimentally observed and quantified?
    - Can these be bypassed and/or corrected?
    - What implications has short-term variability on the use of MBES-measured BS for monitoring on a longer term? Can it hinder the detection of real seafloor changes by the backscatter measurement proxy-approach?

These research questions focused on the reliability/utility of backscatter field measurements by ship-borne MBES for the monitoring of the seafloor interface. Short-term backscatter fluctuations, specifically associated with tidally-induced half-diel variations of the environment, were studied. Field experiments were purposely designed and implemented to acquire repeat MBES data, complemented with a variety of ground-truth data for the interpretation of the backscatter measurements. Three distinct sedimentary and hydrodynamic areas (predominantly: gravelly/clear water, sandy/clear water and muddy/turbid water respectively), reflecting the main surveying environmental settings in the BPNS, were targeted, enhancing the interpretation of the results in the context of longer-term monitoring in these three kinds of environments. The experimental design allowed identifying external sources of variance and their magnitudes, as well as intrinsic seafloor properties responsible for the variability of the measurements. The experimental results are discussed in light of the potential implications of short-term cyclicity on longer-term monitoring.

- **Chapter 5**, related to overall objective C, and entitled “Seafloor change detection using multibeam echosounder backscatter: case study on the Belgian Part of the North Sea” and is a published paper driven by the following research questions:

    - Given both spatio-temporally stable and comparable and directly incomparable serial backscatter measurements, which change detection methodologies to employ?
    - Which kind of spatio-temporal patterns can be observed/quantified?

Detecting seafloor change patterns remains amongst the most challenging tasks in the realm of maritime resource management due to a lack of methodological frameworks, datasets and compounded by issues associated with sensor and data calibration. Therefore, this investigation is based on a serial MBES dataset acquired by former- and current-generation echosounders, combined with a limited number of ground-truth samples, to investigate the suitability of change detection approaches for the identification of temporal trends and patterns of change in seafloor substrate type. An ecologically noteworthy gravel area, situated in the far-field of a concession site of marine aggregate extraction, was targeted.

- **Chapter 6:** Discussion

This last chapter entails a synthesis of the results of this doctoral thesis and discusses the achievements and limitations of the research investigations of Chapters 3, 4 and 5 in the framework of setting up mapping and monitoring applications based on MBES backscatter, bathymetry and ground truth data and in the framework of the MSFD. Future directions and conclusive remarks are addressed herein.

- **Chapter 7:** Conclusions

Conclusive statements.

Chapter 2

2. Multibeam echosounding:
state of the art of hydroacoustic remote sensing

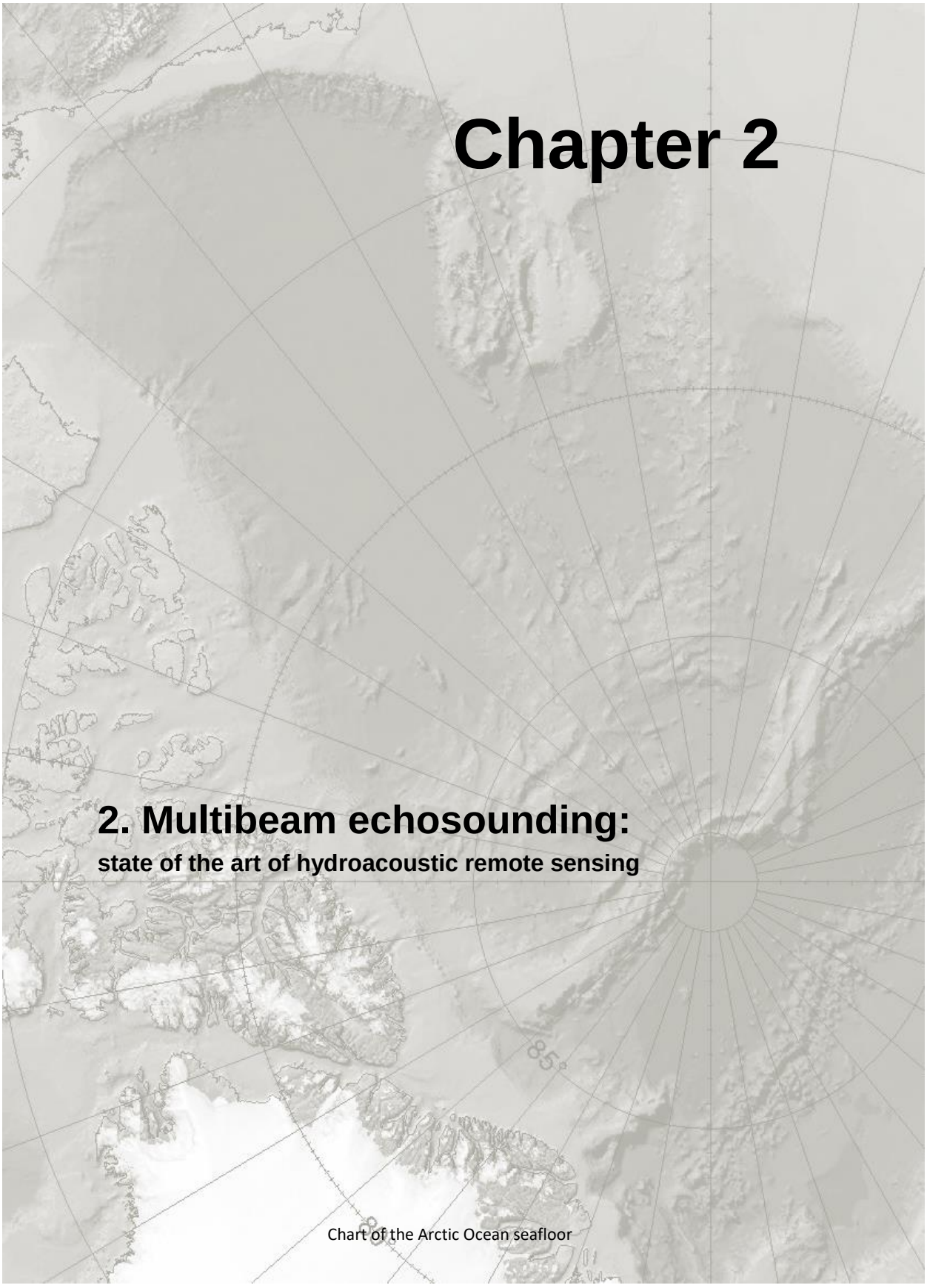


Chart of the Arctic Ocean seafloor

Abstract

A multibeam echosounder system is a state-of-the-art seafloor mapping technology used to co-register continuous bathymetry and backscatter data over broad spatial scales and at a high resolution. Due to this dual data-acquisition character, multibeam sonar technology led to significant improvements in seafloor mapping, compared to the previous generation of map production. The latter was mostly limited to interpolation of ground-truth points and resulted in a considerably coarser depiction of the seafloor sedimentological and geomorphological continuum. This chapter reviews multibeam technology and the associated data types, including operational aspects. An introduction to the theory of acoustical backscattering from the seafloor follows, including an extensive introduction to signal and image-based backscatter data processing. In turn, the modern-day overall seafloor mapping pipeline is described. It focuses on signal and image-based acoustic seafloor classification approaches and the validation of assumptions developed during the analysis of acoustic data by means of ground-truth data collection, interpretation and processing. Key aspects of using backscatter data for mapping and for monitoring conclude this introductory chapter.

2. Multibeam echosounding: state of the art of hydroacoustic remote sensing

If otherwise stated, the technical notions regarding multibeam operation and principles of backscatter in the following chapter are based on Lurton (2010) “An introduction to underwater acoustics: principles and applications”; Lurton and Lamarche (2015) “Backscatter measurements by seafloor mapping sonars: guidelines and recommendations” and the SeaBeam sonar theory of operation manual (SeaBeam Instruments, 2000).

Functionally, a multibeam echosounder system is the maritime equivalent of terrestrial, air- and space-borne remote sensing instruments such as a Synthetic Aperture RADAR (SAR - Radio detection and ranging) and LIDAR (Light detection and ranging) instrumentation (Figure 2.1).

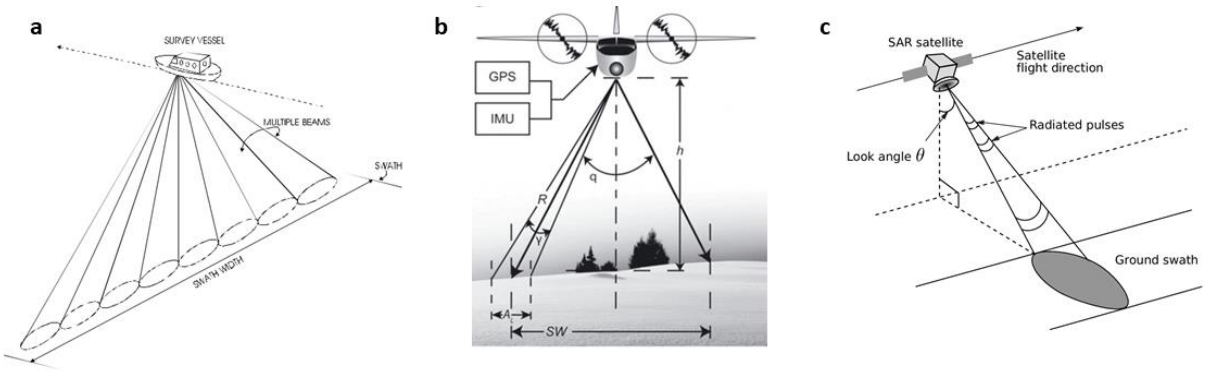


Figure 2.1 - Schematic representation of the observation geometries and similar remote sensing principles for some maritime and terrestrial instruments employed to scan the Earth ‘s seafloor and surface: a) Ship-borne, hull-mounted MBES system, b) Plane-borne LIDAR and c) Satellite-borne SAR.

The physical phenomenon exploited by a SONAR (Sound navigation and ranging) is the mechanical perturbation that constitutes a sound wave (having optimal transmission characteristics in water – Lurton, 2010) whereas terrestrial remote sensing instruments are based on electromagnetic waves and are left beyond their scope for most oceanic environments because of the fast decay of electromagnetic energy in seawater (Kutser et al. 2006). Multibeam systems are *active* sonars: they transmit and receive sound as oppositely to *passive* sonars which only receive sound from secondary emitting sources; e.g. hydrophones picking up the signal emitted by a cetacean, a seismic event or a submarine’s engine noise. Multibeam sonars are designed to transmit and receive several acoustic signals (*beams*) in the millisecond range, which are emitted at a certain *ping rate* through a spatial-filtering process named *beamforming.* These instruments are designed to co-register the two-way travel time it takes for the acoustic pulse to reach and return from a given target (the two-way travel time giving the range and later associated with angle and transformed to depth and bathymetry) as well as the amplitude of the returning echo

(the amount of acoustic energy scattered back to the source of emission – namely the Backscatter Strength (BS)). Because of this duality, MBES systems have become the mainstream tools in seafloor mapping, superseding the former generation of echosounding and seafloor-imaging technology (namely Single Beam Echosounders [SBES] and Side Scan Sonars [SSS] (Kenny, 2003; Brown and Blondel, 2009).

The main units to be considered in understanding the basic functioning of an echosounding instrument are: 1) the transmitter, 2) the receiver, 3) the transducer(s) and 4) the control and display unit. A surveyor programs the control and display unit (i.e. the hardware and monitor of the acquisition software) to order the transmitter unit to *ping*: i.e. to emit a sound pulse. The sound pulse starts by the generation of a controlled electric oscillation at a given frequency and duration. The electric signal is then amplified and converted by the transducer into a pressure oscillation. The pressure oscillation is projected and propagates into the water medium until it returns (echoes) to the transducer as it "bounces back" from the seafloor interface where it gets "contrasted". Upon arrival, the transducer acts as a hydrophone, and reconverts the pressure oscillation to an electrical signal which in turn gets amplified and filtered out of noise by the receiver where the range is estimated based on the timings of transmission and reception of the sound pulse. This information (and possibly the complete return signal) is logged and displayed by the control and display unit. To continuously record this information during navigation, the operation is cyclically repeated at a given *ping rate*, taken as fast as possible while avoiding confusion between adjacent signals. Figure 2.2 provides a schematic representation of this process and these elementary echosounder components (SeaBeam Instruments, 2000).

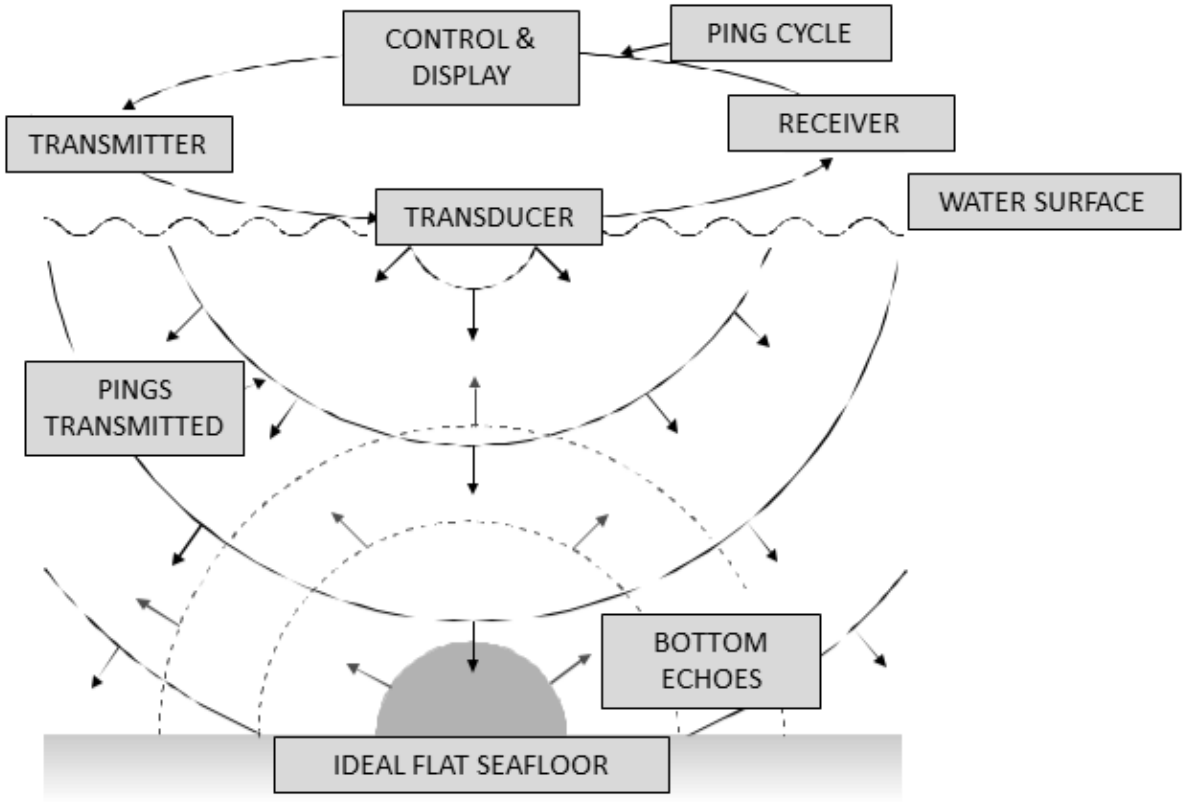


Figure 2.2 – Schematic representation of a single beam echosounder elementary units and the 2D isotropic propagation of a sound wave. A full ping cycle is displayed composed of transmission (black arrows and ellipses), digital-analog (DA) conversion, propagation in seawater and scattering from the seafloor (grey), reception and AD reconversion, amplification and recording. (Readapted from https://www3.mbari.org/data/mbsystem/sonarfunction/SeaBeamMultibeamTheoryOperation.pdf).

The ping emitted by the basic system depicted in Fig. 2.2 spherically spreads uniformly in all directions (i.e. isotropically). Due to this inherent property of sound waves, and the early design of SBES systems, establishing the precise angular direction of the returning echo becomes difficult as the first returning echo may not originate from the seafloor directly below (at *nadir)* the sonar carrier platform given the possible presence of slopes and complex terrains or obstacles. To improve this situation and increase the accuracy of the depth measurement, *directivity* is introduced by focusing the bulk of the emitted acoustic energy inside a narrow solid angle, forming a *beam*. Reflecting the etymology of its name, a multibeam system performs this operation over a higher number of beams at every ping emitted, thus multiplying the number of instantaneous sounding measurements and considerably improving the cost-efficiency of the surveying effort.

Considering two projectors, each emitting identical sound waves spherically spreading through their propagation paths, interference will be established at all points of the propagation medium (Fig. 2.3A). These interferences result as constructive when the pressure peaks coincide and destructive in the opposite situation. The larger amount of energy emitted by two projectors propagates along the directions of

constructive interference, and especially perpendicularly to the separation axis of the projectors creating a spatial selectivity (or *directivity pattern*). Increasing the number of transmitters within an *ad hoc* geometry will concentrate most of the emitted intensity into one main direction named the directivity main lobe (Fig.2.3B). Therefore, targets ensonified in this direction result in stronger intensity in respect to targets ensonified at other directions (i.e. oblique and far grazing angles). The selective projection of acoustic energy allows to control the directivity of the emission; the same principle applies in reception.

A typical projector array configuration is formed by a number of discrete elements, or more simply by one element of sufficient size compared to the wavelength, creating a complex beam pattern. The main lobe corresponds to the maximal energy of the directivity pattern and the projection peak relates to the central part of this lobe. The *beam width* is the solid angle of the main lobe and defines the spatial (angular) selectivity of the sonar. *Sidelobes* are observed on each side of the main lobe; they result from the partial constructive interferences and ideally have to be as low-level as possible (Fig. 2.3B).

This fundamental directivity phenomenon can be generalized. Using discrete elements with appropriate phase or time delays applied electronically (and digitally controlled), it is possible to emit/receive sound waves whose amplitude varies as a function of the angular position through a process named *beamforming,* producing multiple narrow beams steered at preferential azimuthal directions (*beam steering*). By employing a specific configuration of a collection of projector arrays (a series of electroacoustic transducers) perpendicularly positioned to the receiver arrays (a configuration typically referred to as a Mills-cross array), it is finally possible to create a resulting beam pattern with a narrow aperture in both the alongtrack and acrosstrack directions (Fig.2.3C). The continuous coverage achieved by a multibeam system in the perpendicular/across-track direction of the survey vessel is referred to as a *swath*: the result of multiple narrow, dynamically focused and electronically steered beams ensonifying a series of areas on the seafloor surface at given *incidence angles* and forming the typical “fan-shape geometry” (Fig. 2.3C). The swath width of typical Multibeam systems covers from 0° up to 70-80° at port and starboard sides of the vessel equating to an overall 140-160°; this angular coverage can be still extended for dual systems (using transducers tilted on both sides). Within this swath, up to 800 narrow beams can be formed (depending on the system in use), hence providing a high number and density of soundings per ping.

As previously mentioned in the SBES example, within each *ensonified area* (or *footprint*) where the *bottom detection* occurs, the multibeam system Central Unit (CU) measures the range from transducer to seafloor by calculating the time elapsed between transmission and reception (the *Two-Way Travel Time -TWTT*) accounting for the local sound speed monitored by both an ancillary *velocimeter* installed in proximity of the transducers and a Sound Velocity Profiler (SVP) deployed periodically from the ship inside the water column. For beams incident at nadir, the returning echo is processed using the signal amplitude peak (reporting either a single peak value or an average) whereas the differential phase of the signal along the receive array is

used for beams formed in the outer parts of the swath (i.e. interferometrically). The temporal interaction of the acoustic pulse within the ensonified area implies that a time-series of amplitude values can be recorded within it, leading to the registration of the "time-samples" so-called *snippets* (Fig. 2.3C). Hence, it is understood that the backscatter measurements recorded in these beams can contain more information and therefore can resolve and be sensitive to objects smaller than the bathymetric footprint (Miller et al. 1997; Lurton, 2010; Innangi et al. 2015). In other words, there are several time samples of backscatter strength within the footprint of oblique beams, whose number increases with angle from nadir.

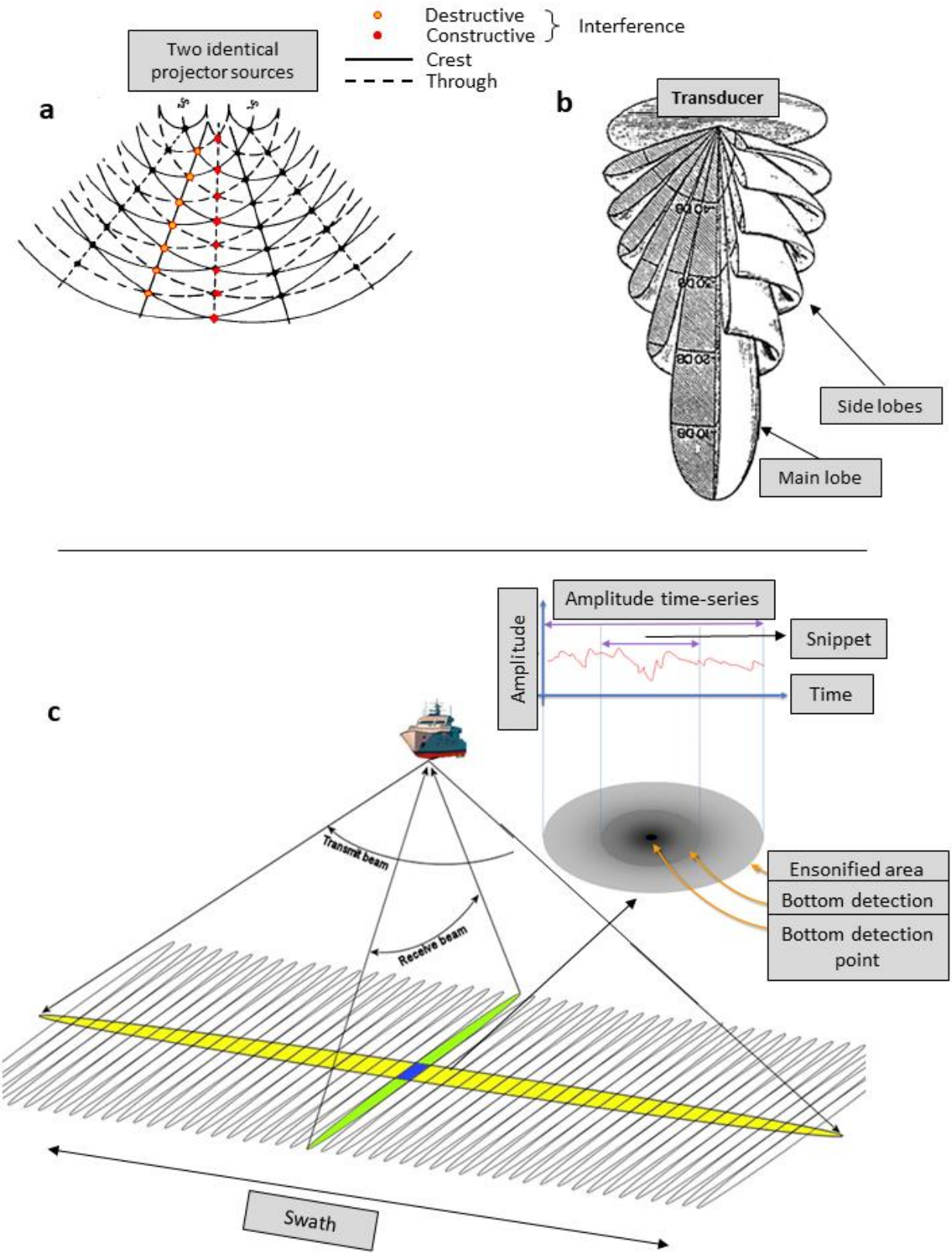


Figure 2.3 – Schematic representation of: a) the establishment of constructive and destructive interferences between two projector arrays emitting identical sound waves, b) the three-dimensional geometry of the beampattern formed by one narrow dynamically focused beam displaying the characteristic main and sidelobes and c) the typical fan-shape geometry of a MBES resulting in the

ensonification of a swath, composed of several ensonified areas. The intersection between transmit and receive beams is shown, indicating the domain where the bathymetric bottom detection occurs as well as the registration of the backscatter snippets as a function of the temporal interaction of the sound pulse within the footprint. (readapted from: a and b) https://qtxasset.com/files/sensorsmag/nodes/1999/838/p28_0399b.gif, c) https://confluence.qps.nl/qinsy/files/en/52101176/52264988/1/1453112810000/Snippet_vs_footprint.jpg

It is important to mention that a series of ancillary sensors are dedicated to compensating the motion of the surveying platform (Fig. 2.4), otherwise adding to the difficulty of retrieving a precisely geo-located position of a returning echo. Briefly, Differential Global Positioning Systems (DGPS) and/or Real Time Kinematic (RTK) sensors are needed for precise geo-location whereas an Inertial Motion Unit (IMU) is necessary to account for the pitch, roll, heave and yaw motions (additional ancillary instruments are referred to in Chapters 3 and 4).

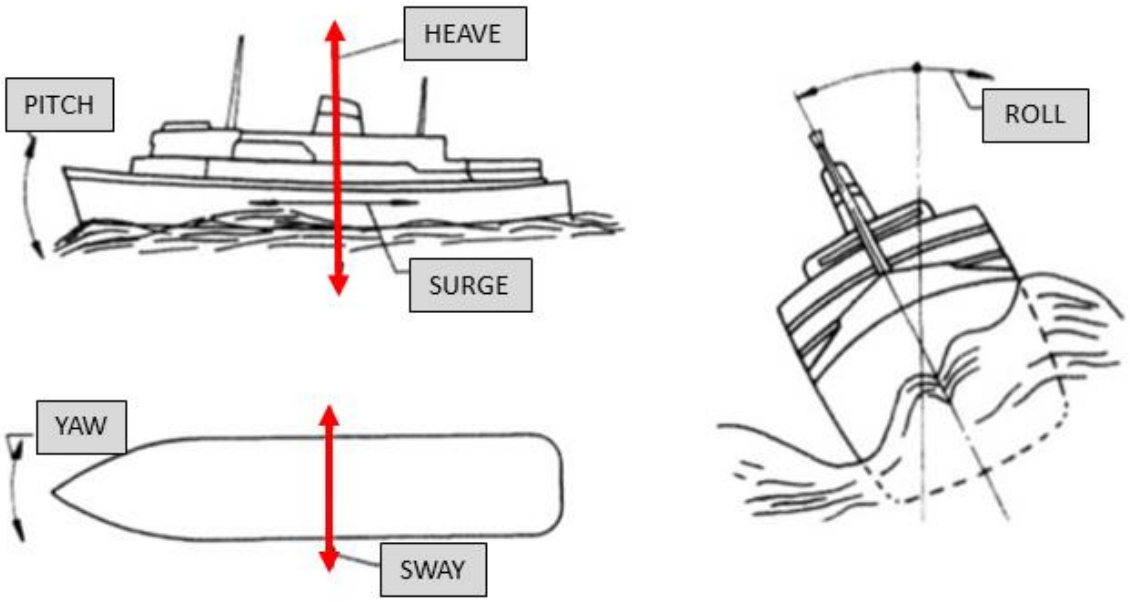


Figure 2.4 – Six degrees of freedom of a carrier platform motion diagrammatically showing pitch, heave, surge, yaw, sway and roll motions. An Inertial Motion Unit and a global positioning system are connected to the multibeam Central Unit to compensate for these motions. A series of tests are dedicated to identifying angular biases between the IMU and the MBES at x (pitch), y (roll)- and z-axis (yaw – heading). Pitch roll and yaw calibration tests are collectively referred to as a "Patch test". (re-adapted from http://generalcargoship.com/ship-motion.jpg)

2.1 Backscattering from the seafloor

A range of factors dictate backscattering of a sound wave incident on the seafloor at a given angle and emitted at a given operating frequency (Jackson et al., 1986; Lurton and Lamarche, 2015). Besides the angle of incidence and the frequency, the backscattering strength is primarily controlled by three fundamental quantities: (1) the *acoustic impedance contrasts* (product of the medium's density and sound velocity)

between the water medium within which the sound wave is propagating and the water-sediment interface onto which it impinges, (2) the surface *roughness* relative to the acoustic wavelength and (3) the *volume* (i.e. in-sediment) inhomogeneity (Lurton, 2010). As such, when an acoustic wave travelling in seawater encounters a medium with a different acoustic impedance (i.e. it collides with an “obstacle”), it in part gets reflected (in the direction opposite to the angle of incidence), in part transmitted (i.e. propagating into the medium) and in part diffusely scattered in all directions. The seafloor thus acts as a new source of emission. The amount of energy that returns to the MBES receiver (i.e. the portion that is backscattered) is referred to as the backscattering strength and its intensity depends on the relative contributions of the mentioned characteristics. Considering a simplified seafloor configuration which is relatively flat and smooth and considering an acoustic signal incident on the seafloor perpendicularly (i.e. at the nadir), the backscatter is then quantified by the reflection coefficient (i.e. the ratio of reflected to incident acoustic pressure) between the different media. In this configuration the incident wave will be reflected opposite to the direction of incidence and will give raise to a specular reflection which is dictated by the acoustic impedance contrasts between the two media (in other words by the “hardness” of the seafloor media).

2.1.1 Acoustic footprint

As previously mentioned, the arrangement of projector and receiver arrays of a MBES system is such that the intersection between transmitted and received beams forms a series of elliptical areas on the seafloor, the ensonified areas (or footprints), together “illuminating” a swath on the seafloor. The footprint A has a circular shape at normal incidence (and results from the intersection of along and across-track beamwidths) whereas it becomes increasingly more elliptic towards the outer parts of the swath (i.e. with increasing slant range). The depth of the area being surveyed, the vessel speed and the angular aperture of the beams together dictate the spatial resolution of the footprint at normal incidence $(\theta \approx 0°)$ whereas for oblique beams $(\theta \neq 0°)$ the ensonified area is dependent on the signal duration (i.e. its pulse length) (Lurton and Lamarche et al. 2015). A MBES system is able to insonify a swat 5 X the water column and at normal incidence (orthogonal to the seafloor surface), is able to resolve fine scale objects, generally measuring on average ~2% of the range for modern MBES systems (Lurton and Lamarche, 2015).

2.1.2 Seafloor roughness scattering

Even in its most regular configurations (such as homogeneous fluid-like sediment), the seafloor is not perfectly flat and presents irregularities at various scales, for example under the form of small-scale bathymetric relief (such as micro-oscillatory sand ripples driven by local hydrodynamics) as well as given by the sediment grains themselves and by the presence of biota and/or other types of seafloor cover (a combination of all is possible given that complexity is the rule rather than the exception in marine sediments – Fig. 2.5). Due to this, part of the wave incident on the seafloor is diffusely

scattered (i.e. reradiated) around the ideal specular reflection. The re-radiation pattern (denoting the scattered field) is dictated by the scale of the surface roughness relative to the acoustic wavelength at a given operating frequency. Practically, a surface that is smooth compared to wavelength causes predominantly coherent specular reflections whereas surfaces that are rough compared to the wavelength scatter the acoustic energy more randomly and homogenously (on the average) in all directions (Fig. 2.5A, B), including back(scattered) to the sonar, greatly reducing the dependence upon the incidence angle. The seafloor roughness texture can be characterised by an array of statistical approaches, including for example the variance of the seafloor relief height, which describes the variation of the surface profile's highs and lows (peaks and throughs) in respect to its average line, and/or by roughness spectral analyses (Ferrini and Flood, 2006; Richardson et al., 2001; Lurton and Lamarche, 2015). It is hence understood that the roughness of the seafloor "perceived" by a sonar system must be considered in respect to the size of the signal wavelength λ. Therefore, depending on the operating frequency, the same seafloor type might result as rough or smooth. Acoustic systems operating at low frequency (i.e. seismic systems used for sub-bottom investigation of the stratigraphic record) results in large wavelengths (in the range 1 – 100 m and operating at frequencies in the range 10 – 3000 Hz), consequently the seafloor is perceived as smooth and its acoustic response interpretation is predominantly based on the specular reflections. On the contrary, for high-frequency MBES systems designed to operate in shallow waters and targeting the mapping of the immediate water-sediment interface, the wavelength is very small (in the range 1 – 10 mm and operating at frequencies in the range 100 – 1000 kHz), hence sensitive to the millimetre scale roughness of individual sediment grains and their geometric configuration (Lurton and Lamarche, 2015). The metric size of the wavelength is bounded to the projected frequency and the speed of sound and it increases with decreasing frequency (Table 2.1). For a 300 kHz emission, such as for the MBES systems used in the present work, and a sound speed in sweater of 1500 ms^{-1}, the wavelength is 0.5 cm, hence being strongly influenced by the sediment grains and offering the potential for surficial sediment characterisation.

Table 2.1 – Typical MBES operational frequency and wavelength relationship at a sound speed in seawater of 1500 ms^{-1}. Indicative values of maximal operative depths for the tabulated frequencies and wavelengths are also given. The operating frequency dictates the wavelength and is therefore a factor of paramount importance to be considered in the interpretation of backscatter data. The same seafloor will be "differently perceived" by the sonar depending on this factor

Frequency (kHz)	Wavelength λ (cm)	Maximal operative depth (m-km)
12	12.5	Deep water (10 km)
100	1.5	Mid-water (1 km)
300	0.5	Shallow water (0.5 km)
450	0.33	Very-shallow water (300 m)

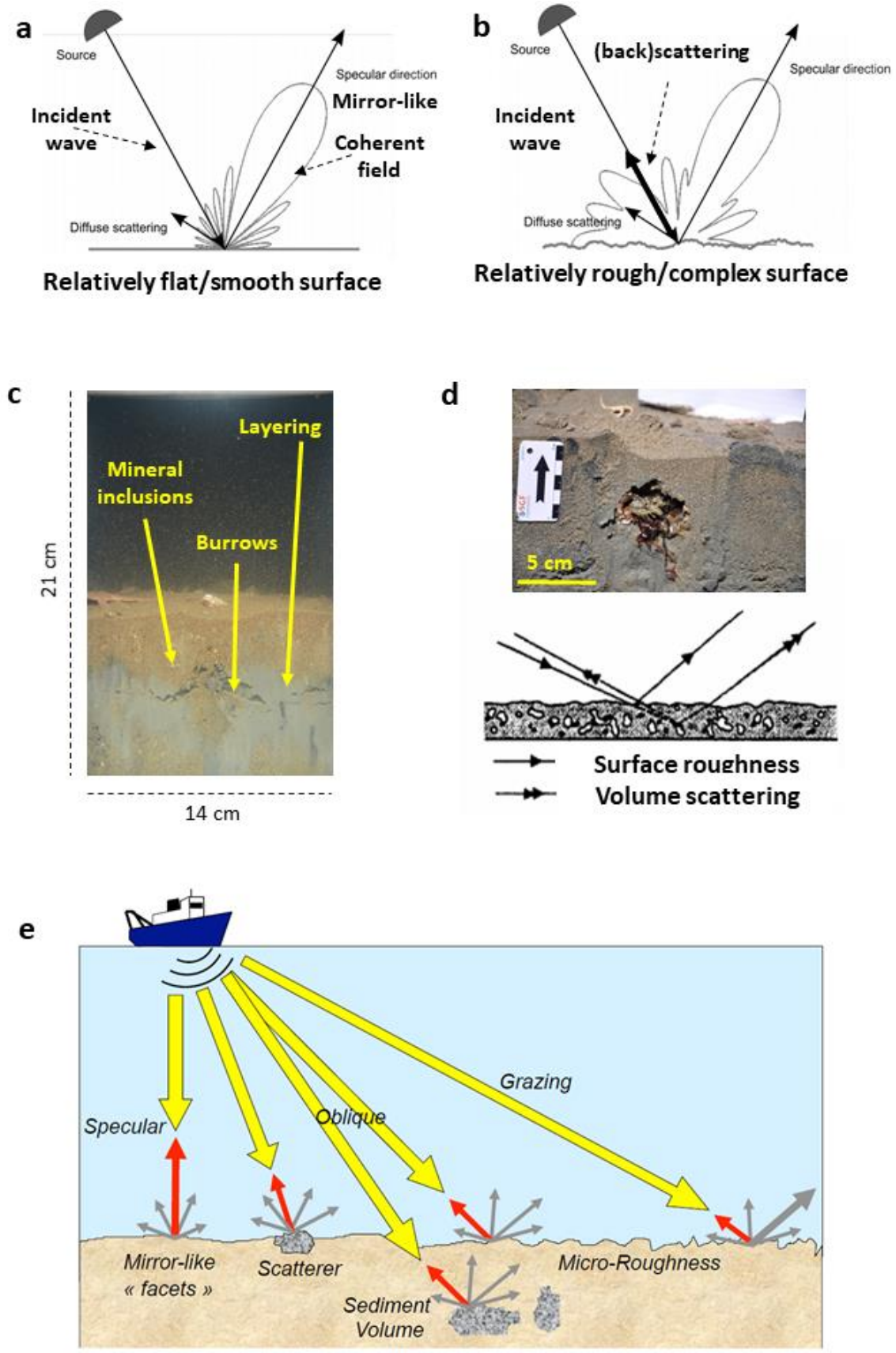


Figure 2.5 – Schematic representation of seafloor scattering mechanisms. Caption continues on next page.

a) Coherent reflection from a relatively flat and smooth seafloor surface, b) Backscattering event from a relatively rough and complex seafloor surface, c) Sediment Profile Imager sample showing the complexity of the sediment surface and inner heterogeneity exhibiting mineral inclusions, infaunal burrows and layering presented as an intercalated matrix of sand and mud. Benthic organisms are also visible at the surface; d) further example of in-sediment heterogeneity from a vertically spit box-core sample. Aside, the volume scattering effect resulting from in-sediment inclusion is displayed. Panel e) displays a summary of reflection, refraction and scattering mechanisms along the echosounder angular line of sight (reproduced from: a and b from https://www.uio.no/studier/emner/matnat/ifi/INF-GEO4310/h13/undervisningsmateriell/sonar_introduction_2013.pdf, d from https://epic.awi.de/id/eprint/26176/1/Bey2006b.pdf and e from http://geohab.org/wp-content/uploads/2013/10/XLurton_Keynote.pdf).

2.1.3 Seafloor volume scattering

Besides the contribution of the seafloor interface roughness to the scattering process, heterogeneities found within the sediment matrix volume can affect the returning echo strength. Volume heterogeneities can be presented as layering of the sediment matrix (Williams et al., 2009), geogenic and biogenic inclusions (mineral and bioclastic detritus such as shell or sponge spicules - Ivakin, 2008), presence of burrows resulting from the behavioural life-traits of certain infaunal organisms (e.g. crustaceans and echinoderms) and presence of gas bubbles (Gorska et al., 2018 - Fig. 2.5C, D): because of the significant impedance contrasts, the latter is the physically prevalent possible cause for volume backscatter. Depending upon the penetration of the incident acoustic pressure into the sediment, such “obstacles” can have a profound effect on the returning echo. The degree of transmission into the sediment depends on the operating frequency (increasing with decreasing frequency), the angle of incidence and the bulk density of the sediment at stake. The greater amount of penetration/transmission into the sediment, the higher are the chances of the sound wave “reaching” buried scatterers and consequently, the greater effect of volume to the backscattering strength: the latter statement applies well when referring to high frequency (100-300 kHz). It has to be kept in mind that at low frequencies (< 100 kHz) the penetration into the sediment increases but the scattering from inclusions decreases, somehow compensating this effect. Transmission into the sediment matrix occurs predominantly at intermediate/oblique angles whereas at nadir, a near-nadir (i.e. steep angles range) the specular reflection dominates (causing potentially obliterating acoustic responses for flat and smooth seafloor surfaces); at far-grazing angles most of the incident energy is reflected, leading to considerably weaker backscattering. Intuitively, the effect of volume backscatter increases for soft and fluid-like sediments. This is because the acoustic impedance contrast between the seawater and a soft and smooth sediment (e.g. unconsolidated mud) is lower compared to a rough and rigid interface (e.g. consolidated rock, densely distributed gravel). Furthermore, soft sediments are prone to host a variety of benthic organisms, resulting in the presence of gas originated from metabolic and/or photosynthetic activity (Gorska et al., 2018).

2.2 Acoustic signals and images

From the above described, it is clear that a very strong link exists between the backscatter phenomenon and the geotechnical characteristics of the sediment type, raising the possibility to characterise these measurements in the interest of geology, sedimentology and biology by the remote sensing proxy approach. In this regard, the backscatter data recorded by multibeam echosounders are considered at two processing levels: (1) angular response (AR) and (2) compensated backscatter imagery (CBI). The following section describes these two kinds of MBES data types.

2.2.1 From angular responses to backscatter imagery

The backscattering strength dependence with angle of incidence on the seafloor is retained as an intrinsic seafloor property directly relating to physical quantities of interest (Jackson et al., 1986; Lamarche et al., 2011; Lurton and Lamarche, 2015). It is a phenomenon of paramount importance in the realm of acoustic seafloor classification as shown in a range of experimental and theoretical studies. Seafloor sediment configurations produce specific angular responses, or signatures, that reflect characteristic properties of the interface roughness, volume heterogeneity and acoustic impedance, raising the possibility to invert (or interpret: i.e. data and or/model driven approaches) the measurements to find links with the seafloor cover at stake (e.g. Daniell et al., 2015; Hasan et al., 2014). Specular reflection and roughness and volume scattering mechanisms reflect the shape and values of the retrieved angular response, forming a “backscatter curve” that resembles the MBES ensonification geometry (Fig. 2.6A, B). The analysis of this kind of information is categorised under signal processing and referred to as angular response or angular range analysis (noted ARA or AR e.g. Fonseca and Mayer, 2007). As previously mentioned, the MBES installation configuration is such that it results in the characteristic fan-shape geometry of ensonification, “illuminating” a wide transversal swath on the seafloor below the carrier platform. The dependence on the angle of incidence results in different scattering mechanisms occurring along this angular range and can be subdivided into three distinct angular domains (depending with seafloor type): (1) the specular region covered by nadir and near nadir beams (typically angles between 0° and 15/30°), (2) the oblique intermediate region (i.e. the oblique range beams between 15/30° and 45/60°), and (3) the fall-off grazing angles region (i.e. the outer range beams between 60° and beyond). The behaviour of the angular response varies with seafloor sediment type: relatively smooth water-sediment interfaces characterised by rather sandy, silty and clayey fluid-like sediments (hence with a relatively low impedance contrast), exhibit strong mirror-like, specular reflections. The strength of the backscatter intensity then quickly decays moving toward the outer portion of the swath. The contribution of volume backscattering is strongest at intermediate oblique angles as it is in this region where the transmission into the sediment is higher for smooth and fluid-like sediments (Fig. 2.6C) while the roughness backscatter contribution is then low. The outer range of the swath generally experiences the weakest returns as most of the incident energy is reflected in the direction opposite to

the incident angle. Contrariwise, water-sediment interfaces characterised by harder materials such as gravel, bioclastic detritus and consolidated rock, tend to yield rather homogeneously distributed acoustic responses (comparatively stable) regardless the angle of ensonification, experiencing a considerable reduction of the specular regime, though still exhibiting a strong decay of the signal towards the outer part of the swath (Fig. 2.6C). The oblique range is generally the most stable area in terms of backscatter strength as the dependence on incidence angle is minimal and the impedance contrast, roughness and volume scattering (the latter increasing for soft sediments) mainly contribute to the intensity of the recorded echo: this is because lesser amounts of seafloor "facets" point back towards the transducer (Lurton and Lamarche, 2015). This angular region is then found as having the highest discrimination potential between different sediment types, forming a stable "plateau" (Fig. 2.6C), and having major implications for seafloor type classification (Jackson et al., 1986; Lamarche et al., 2011; Fezzani and Berger, 2018). The specular region is prone to produce obliterating echoes, considerably reducing the distinction between different seafloor types at these angles (Fig. 2.6B).

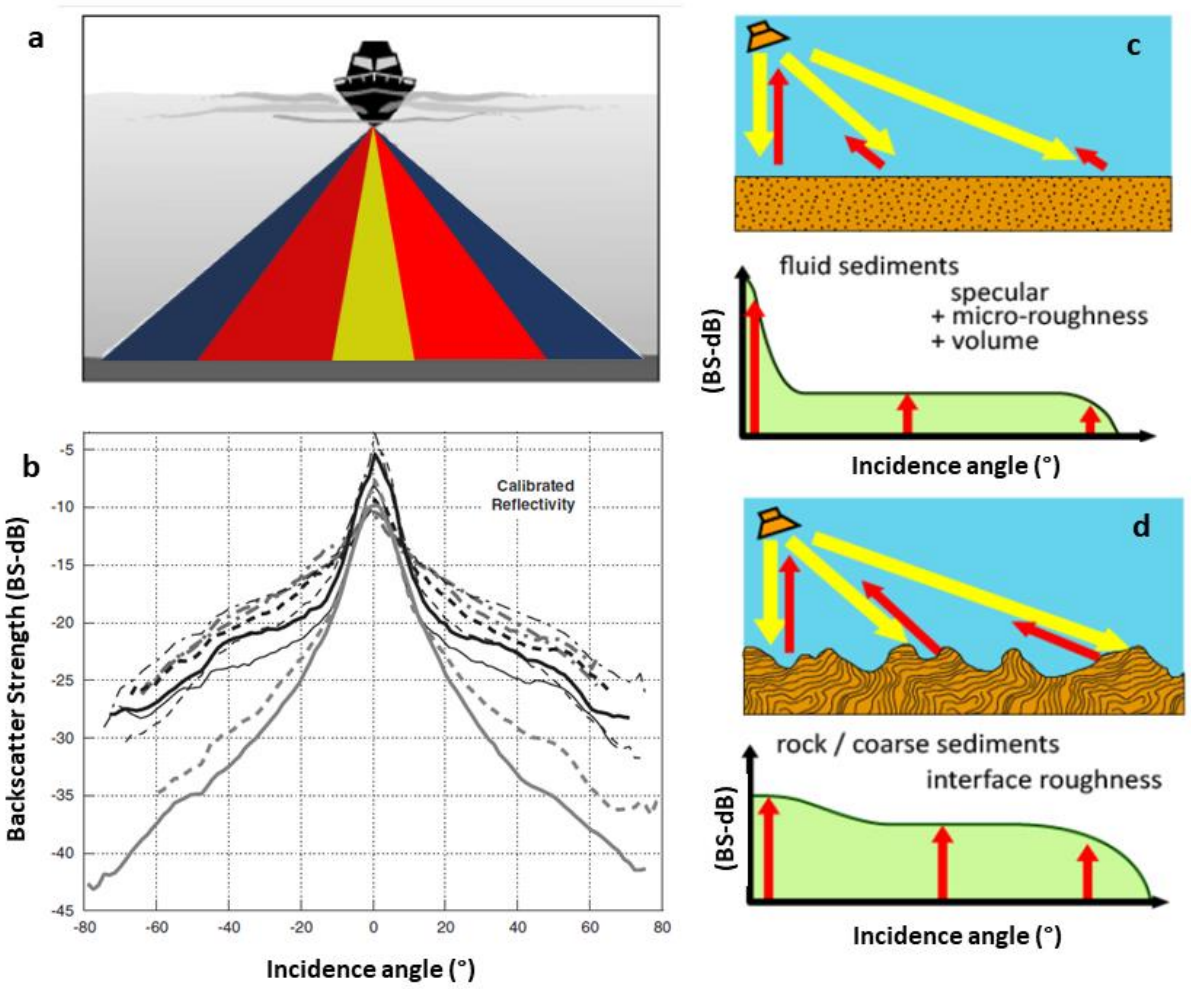


Figure 2.6 – Schematic representation of the angular dependence of seafloor backscatter. Caption continues on next page.

a) fan-shape geometry of a hull-mounted MBES showing the three angular domains with the specular (yellow), oblique (red) and fall-off grazing angles (blue). In b) the backscatter strength variation with angle of incidence is shown for a series of seafloor types (adapted from Lamarche et al. 2011). Panels c) and d) describe the behaviour of the angular response from two ideal sediment configurations: c) a flat soft and fluid-like water-sediment interface with a strong specular reflection and the predominant effect of the micro-roughness given by the sediment grains and by the volume and d) a rocky/very rough water-sediment interface where the acoustic response is primarily driven by the interface roughness (reproduced from Lurton and Lamarche, 2015 (a, c, d) and Lamarche et al. 2010 (b)).

While the angular dependence of backscatter is a keystone feature in the task of acoustic sediment classification and an array of modelling approaches have been formulated on this basis (see section Classification of MBES backscatter), it severely hinders the visual interpretation of backscatter imagery (i.e. geographically gridded format of backscatter data equivalent to the Digital Terrain Model [DTM] of the bathymetry, usually coded in a greyscale) and the application of image-analysis algorithms. As such, statistical compensations are hence required to “flatten” the angular response and produce images in a manner that the seafloor would be observed from a unique fixed angle (Schimel et al., 2018). This is generally obtained by normalizing the data and referencing it to a conventional angle (or a limited range of angles). Typically, the best results of this method can be obtained in the oblique region, around 45°, where the angular dependence is weakest and where the sediment response dominates (Lamarche et al., 2011). A lack of compensation would result in images with a strong banding artefact, displaying higher and lower values at the nadir and oblique angular ranges respectively (Fig. 2.7).

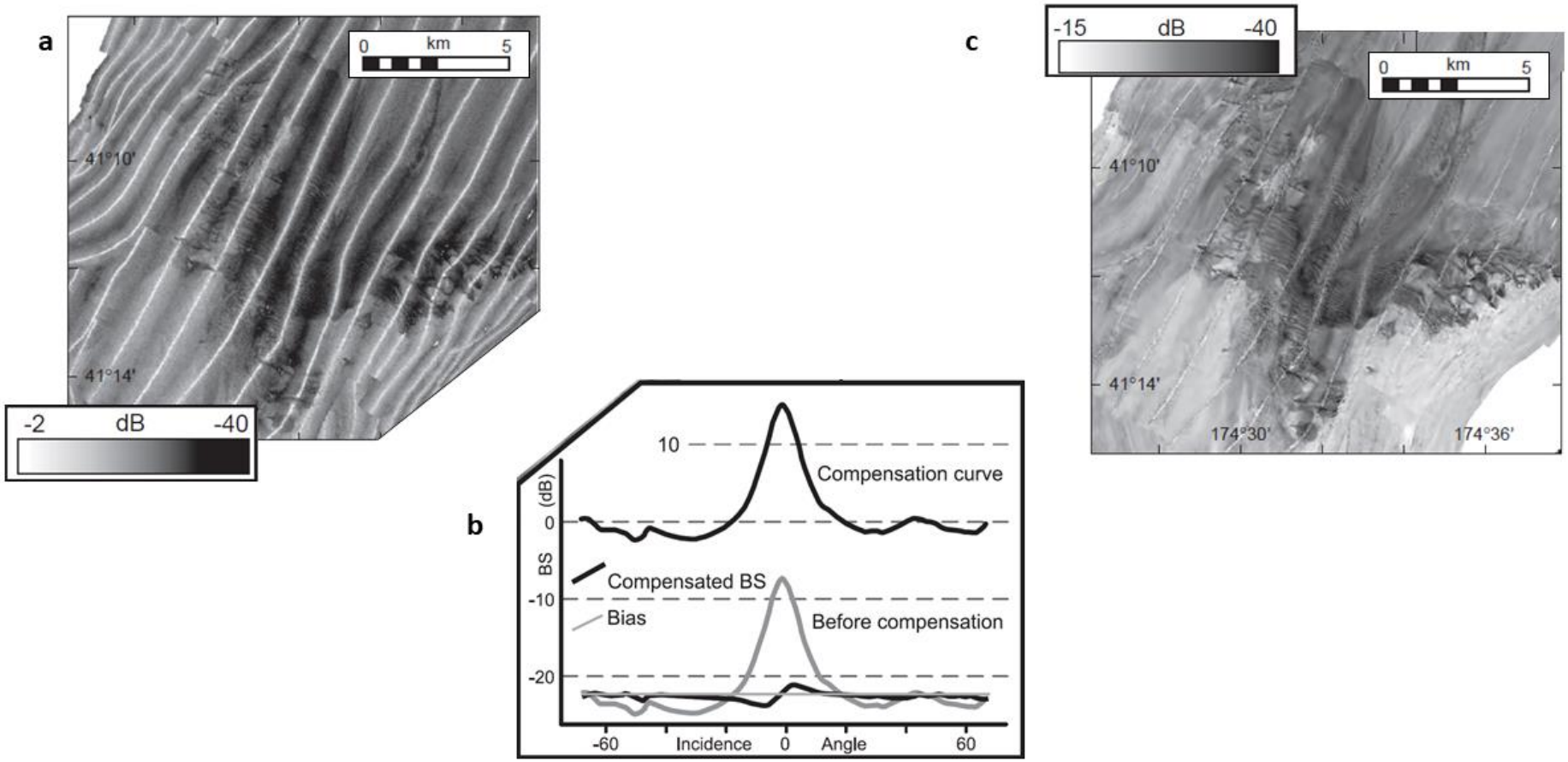


Figure 2.7 – Example of uncompensated and compensated backscatter girds of backscatter data recorded at 38 kHz by a Kongsberg Maritime EM3000 MBES: a) uncompensated backscatter imagery displaying the banding artefact due to the angular dependence. This is particularity visible at nadir and near-nadir incidence angles. B) Gain function used to compensate the backscatter imagery by subtracting a bias value (thin grey line) to the raw backscatter curve (solid grey line). C) Backscatter processed using the compensation curve in b and gridded as a function of the bathymetric pixel resolution. Note how the continuity of acoustic/geologic facies is now observable and interpretable. (adapted from Lamarche et al. 2011) .

The fundamental differences between the two formats of backscatter data (AR and CBI) are the spatial resolution and the type of information they contain. The backscatter AR requires averaging of a series of consecutive pings and processing them over the swath extent or over Regions of Interest (ROIs). The resolution hence approximates the extent of the stack of pings at port and starboard side of the carrier platform. Contrariwise, the CBI can be gridded as a function of the bathymetric resolution and hence it has a considerably finer grain. Nonetheless, due to the compensation of the angular dependence, the mosaicking process leads to a loss of quantitative/physical information, making ground truthing (see *Ground-truth data acquisition and processing*) critical for effective relation to seafloor properties.

2.2.2 Processing and correcting MBES backscatter

As for any remote sensing instrument, the data recorded must undergo a chain of corrections in the processing pipeline to retrieve only the targeted information: in this case the echo information relating exclusively to the seafloor. The backscatter data are inherently noisy, undergoing strong amplitude variations due to the very nature of the scattering process (Lurton, 2010), hence a dedicated data reduction scheme comprising a set of *geometric* (e.g. seafloor topography effect on the actual ensonified area) and *radiometric* (hydrological conditions affecting sound absorption and estimation of the true angle of incidence) corrections must be applied (Beaudoin et al., 2002; Lurton and Lamarche, 2015; Schimel et al., 2018). Over the past decade, dedicated processing platforms have been considerably improved, providing suites of built-in automated routines facilitating the fundamental backscatter corrections which are briefly described hereafter. These operations entail the correction of each beam for a set of terms in the *active sonar equation* (Lurton, 2010). The sonar equation quantifies the performance of a sonar system with respect to Signal-to-Noise Ratio (SNR) by detailing the various components and physical phenomena involved in the sonar operation for a given operation. The terms in the sonar equation are expressed in decibels (dB), referenced to pressure units (i.e. ten times the base 10 logarithm of a ratio of two powers, such as emitted and received intensity). Very classical in all fields of acoustics, decibels are conveniently used due to the huge linear dynamic range of acoustical waves quantities such as sound pressure (e.g. measured in micro pascals - µPa) (Lurton, 2010).

Artefacts that require filtering are attributable to parameters of the recording instrument itself, the seafloor topography and the properties of the medium within which the sound wave propagates (Hellequin et al., 2003). The Backscatter Strength (BS) at a given angle (θ) is primarily accessible from the received Echo Level (EL). This level is controlled by the two-way Transmission Losses (TL) which are in turn driven by the propagation range (R) in what the acoustic pulse spherically spreads (according to $40 \log R$) and by the dissipative nature of the seawater and its *hydrological status* for which empirical absorption coefficients (i.e. α_w, α_v and α_s for seawater, viscosity and scattering, respectively) modulated by temperature, hydrostatic pressure, salinity and turbidity gradients, have been formulated for long and are very generally accepted

(Francois and Garrison, 1982a, 1982b; Urick, 1948; Richards et al., 1996). Furthermore, EL depends on the *hardware characteristics* of the sonar instrument in use; the Source Level (SL) of the transmission, the acoustic signal duration (T), the Directivity Function (D_I) of the transmitted (D_T) and received beam patterns (D_R), the receiver gain (G_R) and the receiver (hydrophone) sensitivity (S_H) and the ensonified footprint ($A\,(R, \theta, T)$). These two "families" of terms (hardware and environment dependent) are reported in the conventional sonar equation (Lurton et al., 2010):

$$EL = SL + D_T - 2TL + 10\log A + BS + S_H + D_R + G_R$$

Important and good practice steps in the data acquisition and processing phases which are user dependent relate to:

(1) Correcting the influence of seafloor topography on the ensonified area using a bathymetric DTM. This allows the backscatter snippets from the *beam time series* (this is a Kongsberg definition of the backscatter snippet registration mode around the central point of the bottom detection – recall from Fig. 2.3) datagrams to be migrated to their true seafloor slope position around the central point of the bottom detection (Fig. 2.3C). This is because the estimation of the seafloor backscatter strength carried out within the CU of the echosounder assumes a flat horizontal seafloor to facilitate computation and georeferencing during data acquisition (Schimel et al., 2018). The bathymetric model itself also requires important corrections regarding the compensation of the tidal oscillation and the removal of spikes. When a survey is conducted, typically consisting in the navigation of a series of parallel track-lines with a certain overlap between them, the tide level harmonically changes resulting in patterns of flood, slack and ebb tidal phases. Ignoring or misestimating this process may result in "stairs-like" artefacts showing "steps" (i.e. offsets) between adjacent track-lines. Therefore, these offsets must be accounted for by the introduction of tide data referenced to a local vertical datum. This is achieved either during acquisition or in the post-processing phase. Tidal information can be obtained from local tide-gauges (for the Belgian Part of the North Sea, reference is made to the Flemish Coast monitoring network; https://sso.meetnetvlaamsebanken.be/) or from RTK GPS logging real-time water level corrections. Further manual and/or automated spike filtering techniques are applied in post-processing to remove outlier soundings.

(2) A second important point is the estimation of the local seawater properties for an adequate computation of the absorption coefficient (α_w) (Chapters 3, 4 and 5) and therefore to adequately estimate the backscatter level (otherwise partly dissipated in the propagation path). Furthermore, knowledge of the sound speed profile is required for an accurate computation of the bathymetry (*ray-tracing*; refraction and bending). For example, in deep water environments (> 1000 m), an improper estimation of the water-column parameters dictating the sound velocity profile would result in inadequate conversions to depth measurements as the sound will refract out of the expected travel-path (Lurton et al., 1994).

Hereafter, the three main sources of unwanted signal fluctuations caused by hydrological conditions and accounted for by empirical coefficients are listed.

Several mechanisms beyond the inherent spherical spreading of the sound wave control the attenuation during the propagation in the seawater medium and can be responsible for unwanted signal fluctuations and noise (Lurton, 2010). Retrieval of the correct target backscatter strength for a given seafloor area must account for the dissipative nature of the seawater medium which inevitably absorbs part of the acoustic energy (Lurton, 2010). Two-way transmission losses expressed in the sonar equation

$$2TL = 40 \log 10(R) + 2\,(\alpha_w + \alpha_v + \alpha_s)\,R$$

result from the contributions of:

- Clear seawater (α_w – in Francois and Garrison, 1982a, 1982b): resulting from absorption in pure water and chemical relaxation processes of boric acid $B(OH)_3$ and magnesium sulphate $Mg\,(SO)_4$. In turn, temperature, salinity, hydrostatic pressure and pH drive the modulation of relaxation frequencies and coefficients.

- Viscosity (α_v – in Urick, 1948) of suspended particles. Viscosity (or inertial friction) results from the density contrast between the particles and the suspending water mass. Due to this contrast, the particle oscillations induced by the sound field will result in a phase-lag in respect to the oscillations in the surrounding fluid. A velocity gradient is established in the boundary layer of the oscillating particle leading to the conversion of acoustic energy to heat and therefore an energy loss in the acoustic field. Nonetheless, thermal absorption can be neglected at sonar operating frequencies (Lurton, 2010). The absorption due to viscosity decreases with increasing grain size as scattering mechanisms take over.

- Scattering (α_s – in Richards et al., 1996). Suspended particles can act as point scatterers leading to the incident acoustic pulse being scattered homogeneously in all directions. These reflections will remain part of the acoustic field, though they will be lost from the transmission/reflection of the sonar pulse. Nonetheless, if part of the energy propagates back to the transducers, the particle scattering can contribute to the increase of volume reverberation (similarly to biotic/pelagic assemblages in fisheries acoustics – Lurton, 2010). Losses due to scattering from suspended particles depend on wavelength (λ) and particle circumference ($2\pi r$ - considering the radius) size and increase with increasing grain size. According to Flammers (1962), when $\lambda >> 2\pi r$ most of the backscattered energy propagates backwards to the transducer whereas in the opposite case, a part of the energy will propagate forward, and the rest diffusely scattered in all directions.

There exists a suite of dedicated oceanographic instrumentation to sample hydrological parameters of the water column and obtain the necessary information for the estimation of TL (Chapter 4 - see for example Fig. 2.8). Various sensors are commonly installed on a Conductivity Temperature Depth (CTD) Rosette platform, i.e. a winch-operated frame which is down-casted from the vessel over the water-column profile. While the CTD samples salinity, temperature and hydrostatic pressure, Optical Backscatter Sensors (OBSs) can be used to estimate the volume concentration of particles in suspension by measuring the backscattered light intensity in respect to the water turbidity (describing the lack of clarity of water). An OBS detects the light scattered from the material in suspension using light sources and photodiode detectors. Laser in-situ Scattering and Transmissometers (LISSTs) are other devices that can be used to sample the grain-size distribution (and parameters of interest, such as the mean particle size) of the material in suspension. They measure the angular intensity variation of light scattered from suspended particles within a laser beam. The nature and concentration of suspended particulate matter (SPM) can also be obtained by the traditional Niskin bottle samplers, also installed on the Rosette platform. These are manually (or automatically) triggered bottles of varying volumes, which once locked, trap the water and sediment in suspension. To determine the SPM concentration, the retrieved samples are filtered, dried and weighted. SPM concentration derived from the water filtration is used to calibrate the OBS sensors. Benthic lander frames (or tripods) are a further type of instrumentation that can be combined with various sensors and deployed at the seafloor for longer times to gather time-series data (Chapter 4).

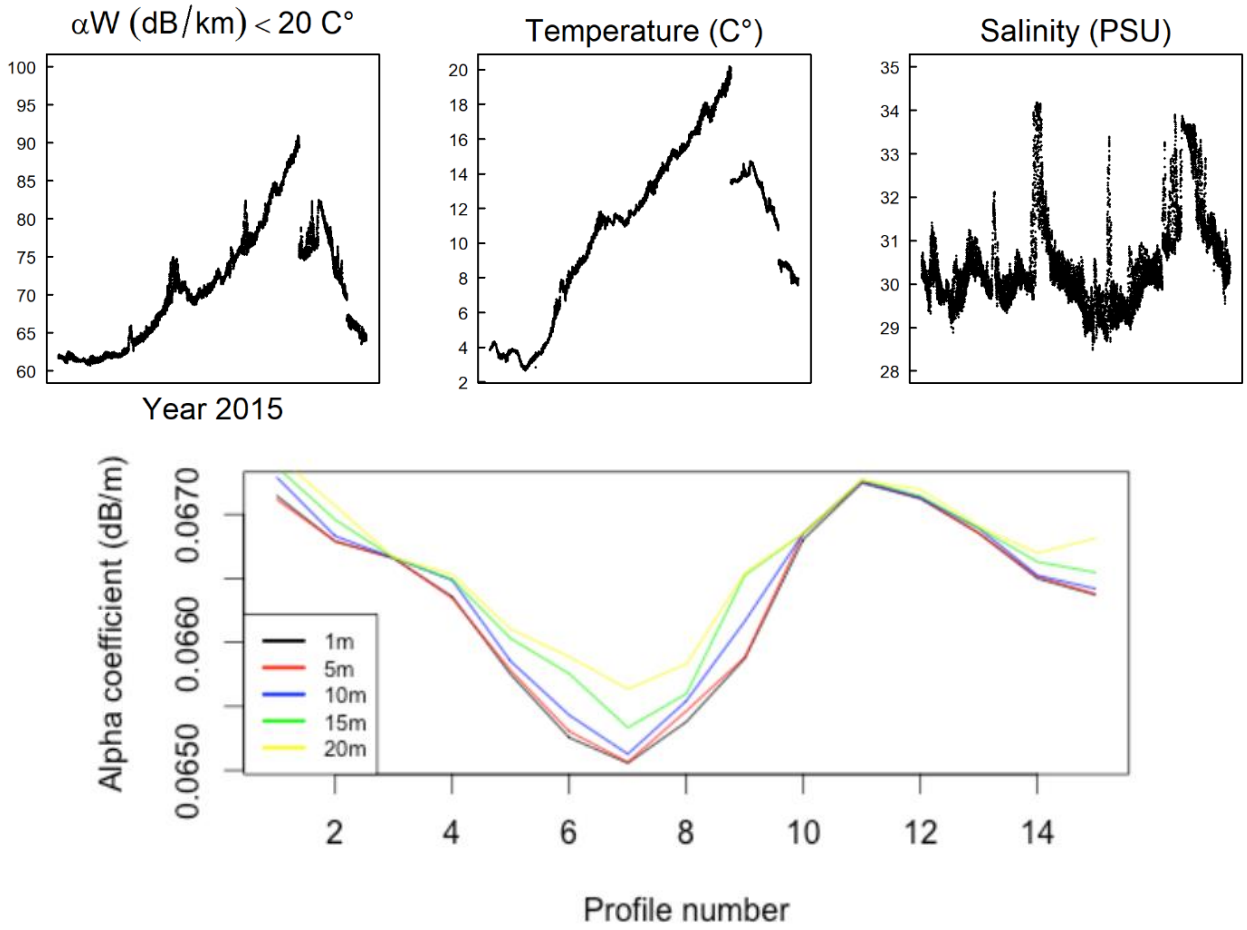


Figure 2.8 – Above: Seasonal (annual; year 2015) variability in seawater absorption with salinity and temperature changes (Data: courtesy of Dr. M. Fettweis, RBINS OD-Nature, Brussels, BE). Below: Example of calculating the water absorption coefficient (T < 20 C°) at different depths over the course of a 13 h tidal cycle in the Kwinte Bank area of the BPNS. The data were acquired from a set of repeated CTD down-casts during the RV Belgica campaign ST1502. Formula used Francois and Garrison, (1982a, b); 300 kHz; Salinity (PSU) and Temperature (C°); pH cst. = 8. (https://odnature.naturalsciences.be/downloads/belgica/campaigns/reports/re2015_02.pdf). Over the duration of the experiment, changes in absorption due to seawater are retained fully negligible (i.e. ~0.002 dB over ~20 m depth and a duration of ~13 h).

2.2.3 Backscatter calibration, repeatability and standards in acquisition and processing

As previously mentioned, the recently increasing interest in using MBES backscatter to explore and monitor the submerged environment, has stimulated the scientific community to develop standards of seafloor backscatter acquisition and processing (similar to reaching the Hydrographic standards). The set of guidelines and recommendations was developed by the Backscatter Working Group (or BSWG; see http://geohab.org/bswg) mandated by the Geological and Biological Marine Habitat Mapping scientific committee (GEOHAB) in 2015. The key aspects of this effort relate to the standardisation of sonar calibration and of best practices in acquisition, processing and interpretation (Lurton and Lamarche, 2015; Schimel et al., 2018) and

estimation of uncertainty (Lucieer et al., 2018; Malik et al., 2018; 2019; Roche et al., 2015, 2018): the common goal being the promotion of a more global comparability of data across echosounder models and acquisition and processing platforms, both in space and time. This is analogous to the terrestrial remote sensing community that benefits from decades of experience (Brown et al., 1993; Fig. 1.2) and constrained by a comparatively less challenging operational environment. Reaching standardisation in the maritime remote sensing community is challenging and is exacerbated by (1) the operational complexity of the surveying environment (Lurton, 2010), (2) ship-based logistics, (3) the number of manufacturers and echosounder models, (4) various dedicated processing platforms, each implementing their own processing algorithms and proprietary software and hardware features (Malik et al., 2018b; Schimel et al., 2018) and (5) technicalities associated with absolute calibration (Eleftherakis et al., 2018).
Absolute backscatter calibration implies reducing the backscatter strength to acoustic quantities that are intrinsic to the target (i.e. reflect physical and inextricable acoustical quantities). It should be perceived as a highly valuable goal targeted by both academics, and manufacturer R&D's (indeed, standardisation of sensors, and related backscatter data, are technological research targets shared across multiple developers, users and companies). The value would come in from the ideal scenario wherein backscatter data acquired by a same echosounder (with standard runtime parameters and a single frequency) system over several oceanographic campaigns, or between different systems and platforms (operating at a given frequency), is directly comparable. However, as an example of the present-day status; data collected in different seafloor areas (or over the same one), by different teams, vessels and systems, and most importantly calibrated or not, will not be comparable.

Absolute calibration carried out by the sonar manufacturer is increasingly advocated to by many (Lurton and Lamarche, 2015; Eleftherakis et al., 2018; Weber et al., 2018), though it remains insufficient. Absolute calibration and standardisation would be especially important in view of ambitious and needed large-scale geographical cover projects such as *The Nippon Foundation-GEBCO Seabed 2030 Project* (https://seabed2030.gebco.net/). This project targets the compilation of a bathymetric chart, though it reflects the enormous potential of working with standardised datasets. A smaller-scale example; a given Exclusive Economic Zone or a marine sanctuary where monitoring is envisaged (e.g. the lagoon of Venice in Italy that was recently mapped in great detail and where careful controls of the data stability and repeatability were carried out – Madricardo et al., 2017), initial mapping operations have the potential to categorising and inventorying n acoustic responses (producing a well-documented library), related to given seafloor characteristics. Serial measurements conducted over such an area (by either the same or multiple systems from different contractors) can then be compared in time, and changes assessed.

Compensating the lack of factory/manufacturer-based absolute calibrations, and the technicalities and facilities required for a rigorous acoustical metrological calibration,

early hybrid/experimental efforts focussing on field-based calibration and inter-calibration propagation between echosounders and vessels (recall from Fig. 1.2), inspired from fisheries acoustics, are recently appearing in literature since the establishment of the BSWG guidelines (e.g. Eleftherakis et al., 2018; Fezzani and Berger, 2018; Ladroit et al., 2018; Roche et al., 2018; Weber et al., 2018). Here, *in-situ* absolute calibration seeks to reduce the measurements to absolute/quantitative values by cross-calibrating the MBES measurements with the equivalent measurements from a calibrated SBES (whether on a stable reference area or not; Eleftherakis et al., 2018; Roche et al., 2018 - Fig. 1.2 and 2.9). The latter is the most pragmatic solution for a hull-mounted system whose deinstallation for a tank-calibration would not be realistic. Alternatively, an *in situ* control of the system's stability and measurements repeatability, requires a natural and well-known reference area (or a fixed/stable target), allowing to relatively compare the serial measurements with a reference survey (a nominal truth) with given acquisition and processing parameters (Chapters 3, 4 and 5). In any case, whether working with uncalibrated (relative calibration) or fully-calibrated system, consistency is the important aspect to consider and it is here that *repeatability* and *standards in acquisition and processing* become necessary. This includes, for example, assessing the *instrumental* and *environmental* drift on the short- to medium- and long-term. For example, such a stable reference area enables controlling the stability and linearity of the measurements of the system in use over the duration of the mapping/monitoring program (Roche et al., 2018 - Fig. 2.9).

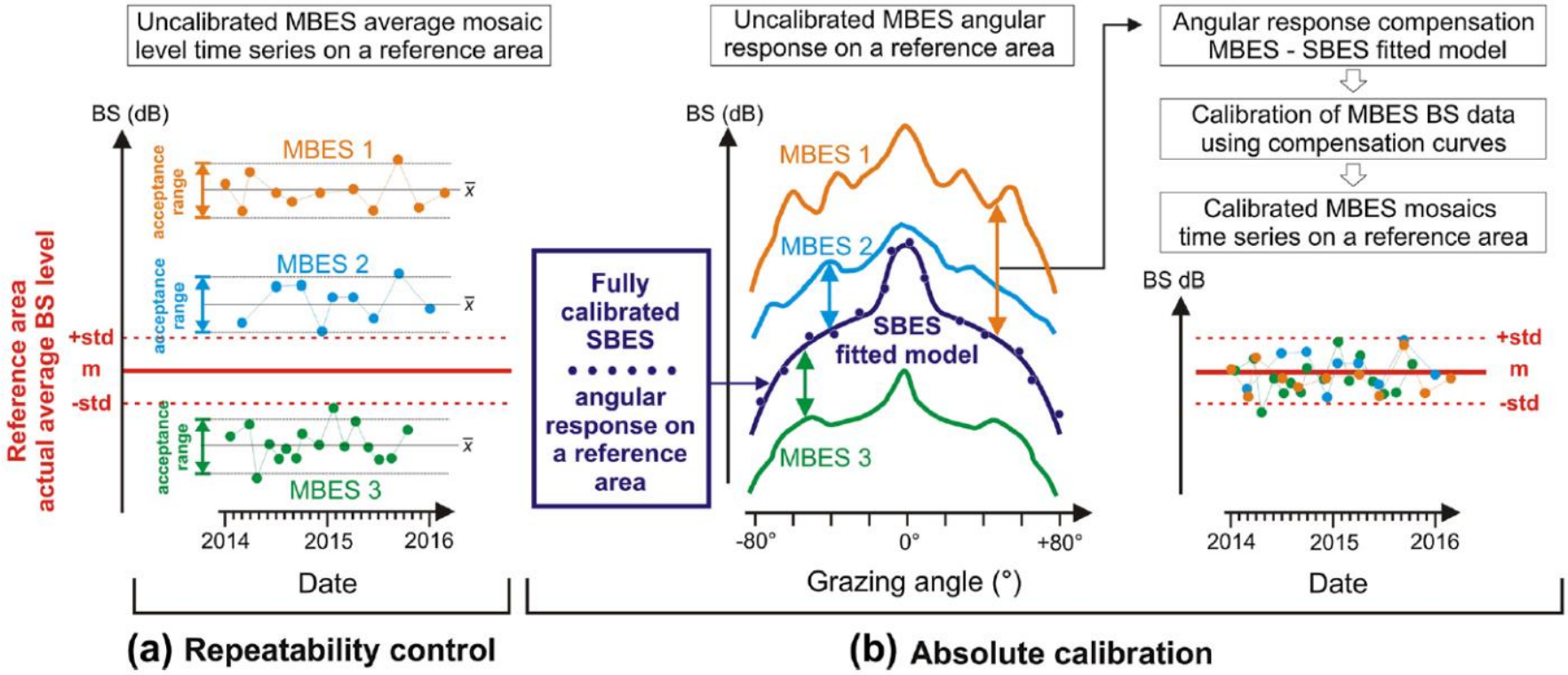


Figure 2.9 – A comparison between assessment of the backscatter measurement repeatability based on a natural reference (a) and the inter-calibration propagation between a strictly calibrated SBES and other echosounders (b). In (a) the stability of a series of repeated measurements on a well-known (in terms of target strength at a given frequency) stable reference area is appraised against a sensitivity threshold (i.e. the ± 1 dB inherent transducer sensitivity reported for Kongsberg systems [Hammerstad, 2000]). In (b), a calibrated SBES angular response model is used to derive the bias (in dB) to be applied to the other measurements acquired over a same area (any area with favourable conditions – see Eleftherakis et al., 2018). The result is the full inter-comparability of backscatter measurements from different echosounders. (image taken from Roche et al., 2018).

In the coming years, it is expected that a global community response to the issue of calibration (triggered by the GEOHAB-BSWG ongoing works) of swath mapping echosounders will promote the provision of factory calibration (absolute calibration conducted both at the Factory Acceptance and Sea Acceptance Test levels) by the manufacturers, largely improving inter-comparability and, for example, making possible future merging of disparate datasets producing large-scale geographical coverage.

2.3 Classification of MBES backscatter

A diversity of approaches to classify (i.e. arrange in groups by property) the remotely-sensed data has been developed over the past few years (e.g. Brown et al., 2011). As previously mentioned, MBES backscatter is presented at two processing levels: AR and CBI data types. Consequently, classification is categorised into signal- and image-based approaches. The following section introduces signal- and image-based approaches dedicated to the prediction of seafloor sediment type. Additionally, it is worth noting that two main categories of seafloor substrate classification exist and can be referred to as 1) model-driven and 2) data-driven. In the first approach, denoting the early and pioneering approach, the logic is that of predicting the seafloor type based on input physical parameters which have been rigorously measured and modelled. Therefore, stringent physical priors constrain these modelling approaches, often neglecting the complexity of the environment. On the contrary, data-driven modelling, such as application of machine learning, reflects the complexity of today's scientific problems, dealing with large and heavy data volumes and seeking to classify the seafloor by enabling algorithms to learn complex structures and generalise/predict unseen patterns. Extensive reviews of ASC have been compiled by Simard and Stepnowski (2007) and more recently by Brown and Blondel (2009) and Brown et al. (2011).

2.3.1 Signal-based approaches

In signal-based classification, the prediction target is the physical dependence of backscatter intensity variation with angle of incidence at a given frequency (e.g. Che Hasan et al., 2014; Daniell et al., 2015; Alevizos and Greinert, 2018; Fezzani and Berger, 2018). Modelling of the angular response has primarily ramified into geophysical (i.e. physical/geoacoustical models – e.g. Jackson et al., 1986) and empirical (i.e. phenomenological/heuristic – e.g. Lamarche et al., 2011) AR models. Physical models target the prediction of the AR "behaviour" based on the tuning of input geo-acoustical parameters, obtained *a priori* from a given area and compare the predictions with calibrated field measurements. Empirical models directly target the statistical fitting of measured angular responses that can in turn be physically interpreted to describe seafloor parameters and type. Two well-established examples are presented hereafter. Overall, backscatter models can be summarised as "Model-driven" (Physical) and "Data driven" (Empirical). The latter kind of modelling should be

perceived as highly advantageous given the complexity of “solving the inversion problem” based on physical/model-driven approaches (Anderson et al. 2007).

2.3.1.1 Physical geoacoustical modelling of backscatter angular response

One example of predicting the angular response backscatter is the composite-roughness model, a part of the complete model known as the Applied Research Laboratory of the University of Washington (APL-UW, 1994 - Jackson et al., 1986). It is the part of the model dealing with the interface roughness scattering, completed by the volume component and replaced by other empirical solutions when studying non-sedimentary seafloor types. The model was developed for low to medium-high frequency backscatter (≤ 100 kHz). Sound reflection and scattering mechanisms (i.e. roughness and volume) are encompassed in this model and parameterised by a set of tuneable parameters resulting in the prediction of a set of generic seafloor sediment types ranging from silt to rock and shown in Figure 2.10 (input geo-acoustical parameters are mentioned therein). This model applies a facet reflection approximation (often referred to as the Kirchhoff approximation) at steep angles (nadir and near-nadir range) and sums together a roughness and volume model at lower incidence angles (oblique and far-range). The model summarises in the sonar equation the target strength (TS) of the bottom (BS) within the ensonified footprint (A):

$$TS = BS + 10log_{10}(\sigma_r + \sigma_v) + 10log_{10}A,$$

where σ_r and σ_v are the backscatter cross-sections of interface roughness and sediment volume respectively (Lurton and Lamarche, 2015). A range of studies have employed such a modelling approach and have demonstrated the connection between sediment grain-size parameters and the backscatter AR (e.g. Hughes-Clarke et al., 1997; Fonseca et al., 2002; Fonseca and Mayer, 2007). Nonetheless, since the model assumes homogenous sediment properties and isotropically distributed surface roughness (i.e. ideally simplified configurations), its applicability to a global scenario is limited given that complexity is a rule in marine sediments rather than an exception, and the model is limited by the need of *a priori* knowledge of the sediment properties for a given study area. Furthermore, this model was developed for frequencies ≤ 100 kHz, making it inapplicable for measurements acquired using contemporary MBES which generally operate above 200 kHz. Finally, the non-sediment seafloors (sandy gravel and coarser) are not consistent with the physical model (dedicated to fluid-like sediments) and are modelled by empirically-established formulas.

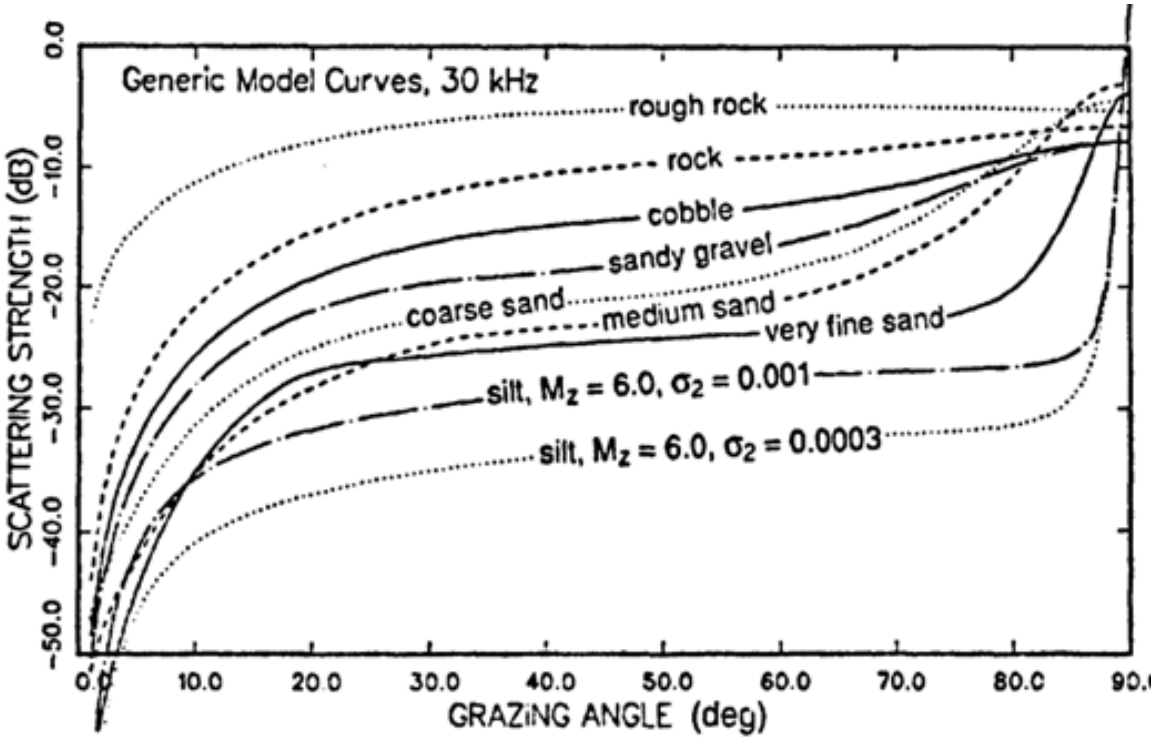


Figure 2.10 – Angular response curves modelled backscatter strength at 30 kHz for a set of generic bottom types. The input parameters tuned to produce these responses are taken from a set of empirically measured and theoretically derived parameters including the bulk grain size, the sediment density ratio, the sound speed ratio (between water-sediment interface and the seawater), a loss parameter, volume parameter and spectral exponent and strength (image taken from: APL-UW Technical Report TR9407, 1994).

2.3.1.2 Empirical modelling of angular responses

A pragmatic approach, bypassing the complexity of the considerable number of parameters required for an effective and globally applicable solution of the “inversion problem”, is to employ phenomenological/empirical modelling approaches (Chapter 3). One such approach is the Generic Seafloor Acoustic Backscatter model (GSAB – Fezzani and Berger, 2018; Lamarche et al., 2011). The model is used to fit field measurements that have been geometrically and radiometrically corrected and with omission of possible manufacturer built-in compensations (following processing procedures introduced in e.g. Fezzani and Berger, 2018 and Roche et al., 2018 - Chapter 3 and 4). Once the field measurements have been retrieved as angular responses, the following step tunes a set of statistical distributions until an optimal fit to the raw data points has been found (black crosses and distributions in Fig. 2.11). A combination of three statistical angle dependency laws are used to capture the angular response: (1) a Gaussian law fits the specular region of the angular range, (2) a Lambert-like law is used to fit the oblique and fall-off regions, and (3) a second Gaussian law fills the intersection between the former components. Overall, four to six parameters (A-D or A-F) are used to describe the angular response in terms of dB intensities and angular extents, thus the behaviour of backscatter as a function of incidence angle for a variety of sediment types. Figure 2.11 provides an example of

fitting GSAB parameters to measured data. The parameters used to fit the measured angular response do not directly relate to geological and geotechnical sediment properties as in the geo-acoustical backscatter models (i.e. Jackson et al., 1986), though they provide a physical description of the backscattering from seafloor sediments, making unsupervised, yet broadly descriptive, inferences of the substrate type possible for any field measurement. However, to establish detailed relations between the modelled angular responses and the sediment properties, ground-truth data are required. The model reads:

$$BS(\theta) = 10\log [A \exp (- \theta^2/2B^2) + C \cos^D \theta + E \exp(-\theta^2/2F^2)]$$

For example, the parameter *A* is related to the specular coherent reflection (the maximal amplitude intensity in dB) and it will be highest for smooth and fluid-like sediments and for strong contrast in acoustic impedances between the water medium and the seafloor interface. *B* refers to the angular extent of the specular domain and relates to the interface roughness. The parameter *C* relates to the mean backscatter level (in dB) in the oblique angular range. This parameter is associated to the Lambert's law which describes the backscattering phenomenon at oblique angles for rough and coarse water-sediment interfaces (at a roughness scale comparable to the acoustical wavelength). *C* is also dependent on the volume, thus on the in-sediment inhomogeneity. It is found to increase with increasing roughness and impedance, as well as in the presence and characteristics of buried scatterers, possibly being the dominant scattering mechanisms in soft sediments. Parameter *D* is the decay rate of the backscatter strength (in dB) with grazing angle and is found to increase for soft and flat seafloor interfaces. Without clear physical meaning, the parameter *E* describes the maximum level (in dB) of the transitory region between the specular and grazing angular domains whereas *F* refers to its angular extent (Lamarche et al. 2011).

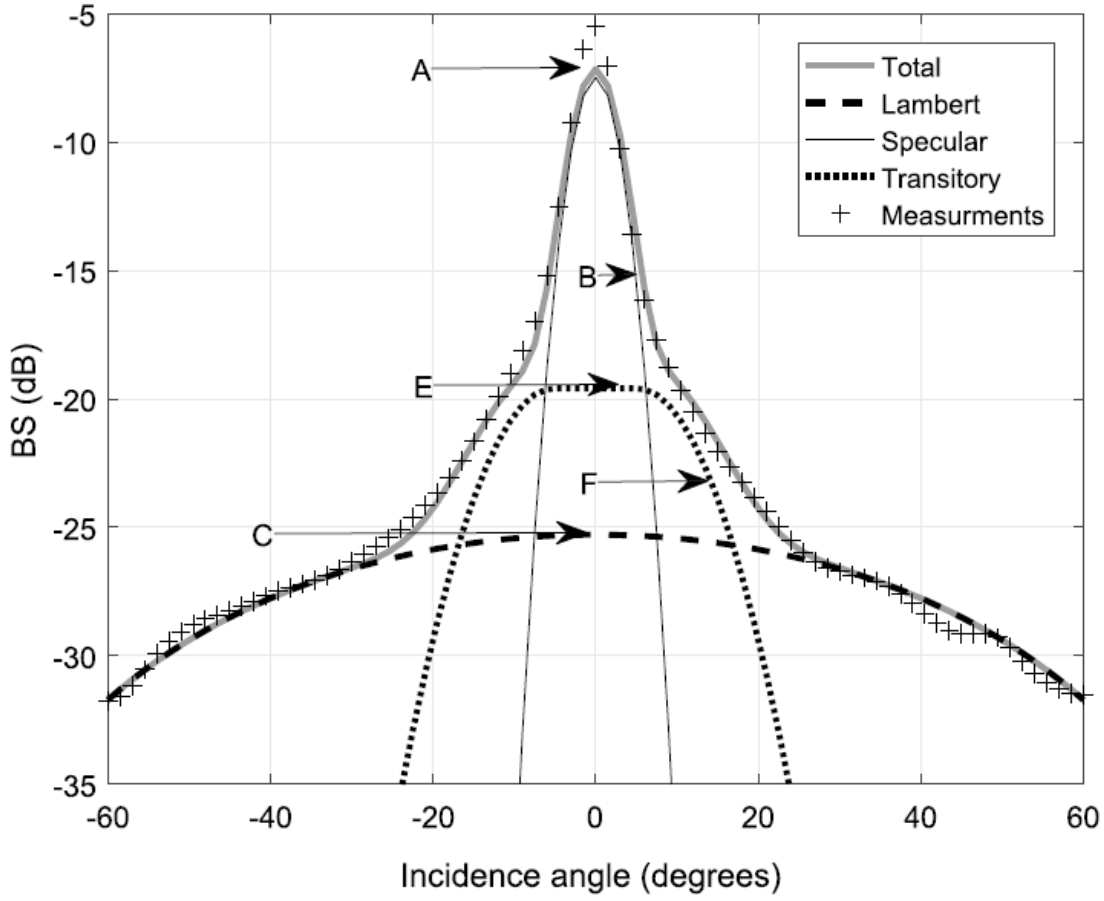


Figure 2.11 - Fitting of GSAB model (grey solid line) to measured and corrected raw backscatter data (black crosses) via a combination of Gaussian and Lambert-like functions (dashed and solid black lines) adjusting parameters A to F. The adjustment of the parameters is carried out iteratively until an optimal fit to the measured data is found. The model will find a symmetric curve (from Fezzani and Berger, 2018).

2.4 Image-based classification

A second and most widely embraced type of backscatter data classification is undertaken using image-analysis. This technique has largely benefitted from the recent design of MBES systems (co-registering depth and backscatter), the improvement of angular dependence compensation algorithms, in particular the reduction of the nadir artefact (i.e. Fonseca and Calder, 2005) (Chapter 3), and from the significant improvements in data integration and visualisation into Geographical Information System (GIS) platforms (e.g. Breman, 2002).

Here, the goal is to segment the data into homogenous units/areas (contiguous sets of pixels) representative of different acoustic facies, i.e. *"the spatial organization of seafloor patches with common acoustic responses and the measurable characteristics of this response"* (from Lurton and Lamarche, 2015). In this context, the data types used relate to a raster-format of compensated backscatter imagery, presented in a gridded georeferenced frame and commonly depicted with a 256 grey level colour palette and to bathymetric surfaces as DTMs. Traditionally, subjective and

expert interpretation of the images was applied by a geologically-oriented manual digitisation of backscatter mosaics. Nonetheless, the preponderance of ever-increasing volumes and multidimensionality of datasets being treated (i.e. comprising hydroacoustic measurements, model results, and substrate and biology ground-truth data), has led to a paradigm shift in the production of thematic maps that currently focus on the adoption of automated, objective and repeatable methods seeking to find mathematical relationships within the multivariate input dataset (Stephens and Diesing, 2014). Over the past decade, the ASC field has placed much attention on the automatization of image classification and on comparative research of the available classifiers (e.g.; Ierodiaconou et al., 2011; Hasan et al., 2012; Calvert et al., 2014; Diesing et al., 2014; Montereale Gavazzi et al., 2016; Gaida et al., 2018). In image-analysis, single pixels define the elementary spatial grid usable for classification as oppositely to AR signatures derived from larger seafloor patches. The high-resolution obtainable in image-processing implies the ability to identify spatial-units down to the level of single patches (Hitt et al., 2011) and single biogenic features such as sponges (Montereale Gavazzi et al., 2016). This has had major benefits in marine ecological studies targeting benthic habitat mapping and seascape ecology (e.g. Kostylev et al., 2001; Galparsoro et al., 2009; Monk et al., 2010; Wedding et al., 2011).

A significant advantage of image-analysis is in the exploitation of a broad spectrum of bathymetric (morphometric derivatives form DTMs) and backscatter (statistics and textural attributes of backscatter) derivatives (Chapters 3 and 5). Derivatives are computed using neighbourhood analysis within windows of varying sizes and shapes (typically a 3 x 3 window, but this will vary on the scale targeted in the analysis). Bathymetric derivatives such as measures of roughness, slope, curvature, aspect, eastness and northness and the Bathymetric Position Index (BPI – Lundblad et al. 2006), hold great potential in seafloor classification as they are found to relate to the geomorphological organisation of the seafloor, the susceptibility of sedimentary and hydrodynamic processes, and to occupancy by benthic biota (McArthur et al., 2010; Harris, 2012). Indeed, since the advent of MBES systems, the field of marine geomorphometry has seen a drastic development in the past few years (Lecours et al., 2016a, 2016b) and has become a discipline in its own right, ramifying from the well-established field of terrestrial geomorphometry (Pike, 2000). A second kind of derivatives are those computed from the backscatter imagery, typically relating to textural attributes of the acoustic energy using Grey level Co-Occurrence Matrices (GLCMs - Haralick and Shanmugam, 1973). Backscatter textural analysis using GLCMs has been extensively used in seafloor mapping using MBES (Blondel et al., 2015; Blondel and Sichi, 2009; Micallef et al., 2012; Montereale Gavazzi et al., 2016; Prampolini et al., 2018) and has shown great potential for substrate characterisation. Textural indices such as entropy, measuring the lack of spatial organisation within a computation window and akin to roughness, and homogeneity, measuring the amount of local dissimilarities (i.e. local organisation), have been argued to suffice in capturing the textures visible in sonar imagery (Blondel, 1996). Additionally, backscatter statistics computed within a kernel, such as mean, mode, median, minimum and

maximum, have been used by a range of authors (e.g. Eleftherakis, 2013; Snellen et al., 2018) to improve the predictive performance of the model in use.

The calculation of derivatives from the primary MBES data leads to the production of a “stack of predictor layers” relating to the explanatory variables in the classification process (also together referred to as the information system). This is commonly treated as a composite image, similarly to land-cover mapping in terrestrial remote sensing applications. The response variable, i.e. the classification target, is then the set of ground-truth samples collected at given locations within the surveyed area. Ground-truth samples are required to estimate the predictive accuracy of the classifiers applied (see Ground-truth data acquisition and processing).

Image-analysis can be subdivided into *unsupervised* and *supervised* classification approaches. These two data-aggregation concepts/approaches are introduced hereafter and schematically represented in Figure 2.12.

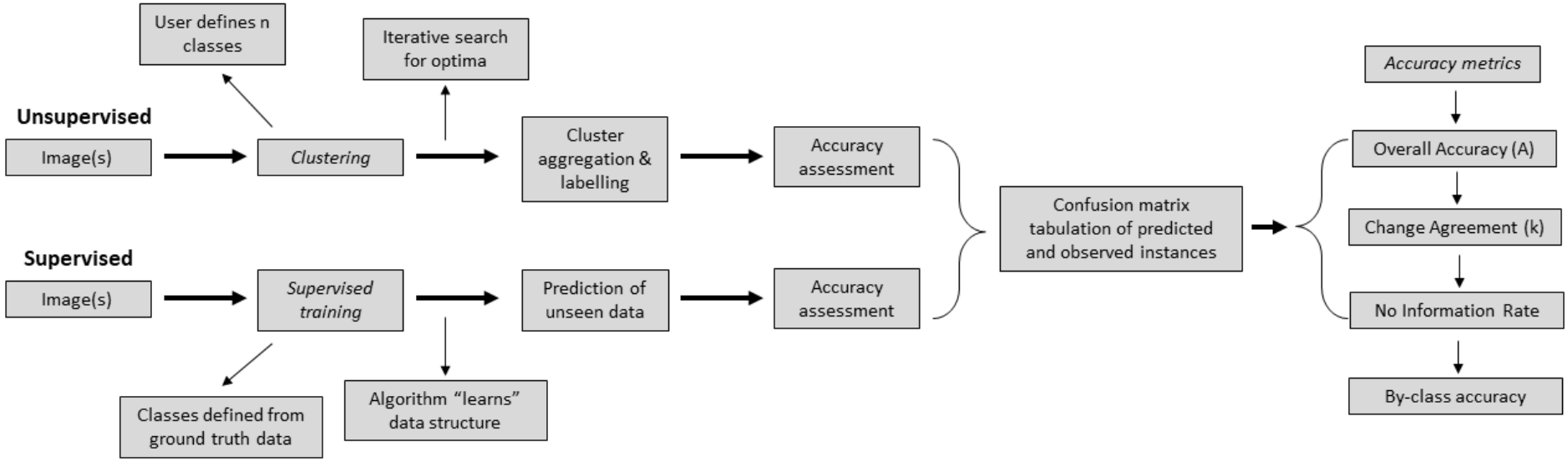


Figure 2.12 – Schematic representation of unsupervised and supervised image classification procedures.

2.4.1 Unsupervised image-analysis

Unsupervised image classification seeks to find statistically inherent natural groupings/patterns in the unclassified/unlabelled data. In this regard, an array of clustering techniques has been used in seafloor mapping literature, particularly those in the family of partitive clustering (Lathrop et al., 2006; Lucieer and Lucieer, 2009; Lucieer and Lamarche, 2011; Eleftherakis, 2013; Calvert et al., 2014; Snellen et al., 2018). A frequently used clustering algorithm is the k-means (Hartigan and Wong, 1979) (Chapters 3 and 5). k-means seeks to reduce the within-cluster variance while maximising the variance between groups through an iterative process of cluster centres assignment and re-allocation (schematically represented in Figure 2.13). The number of classes (i.e. the optimal partition of the observations into k clusters) is either user-defined based on expert interpretation or searched using cluster-validation criteria such as the Within Group sum of Squared Distances (WGSSD) or the Silhouette coefficient (Rousseeuw, 1987) (Chapters 3 and 5). The former is a measure of cluster homogeneity and looks at the WGSSD as a function of the number of clusters. The optimum is chosen where adding clusters does not improve the WGSSD. The latter quantifies the dissimilarity of single data points to the overall points of its cluster and returns measures in the range 0-1. A Silhouette coefficient > 0.5 is indicative of sufficient class separation whereas below this threshold classes are found to be overlapping (Eleftherakis, 2013). Once the optimal number of classes has been estimated, clustering is applied, and the following steps are to assign labels to the identified groups.

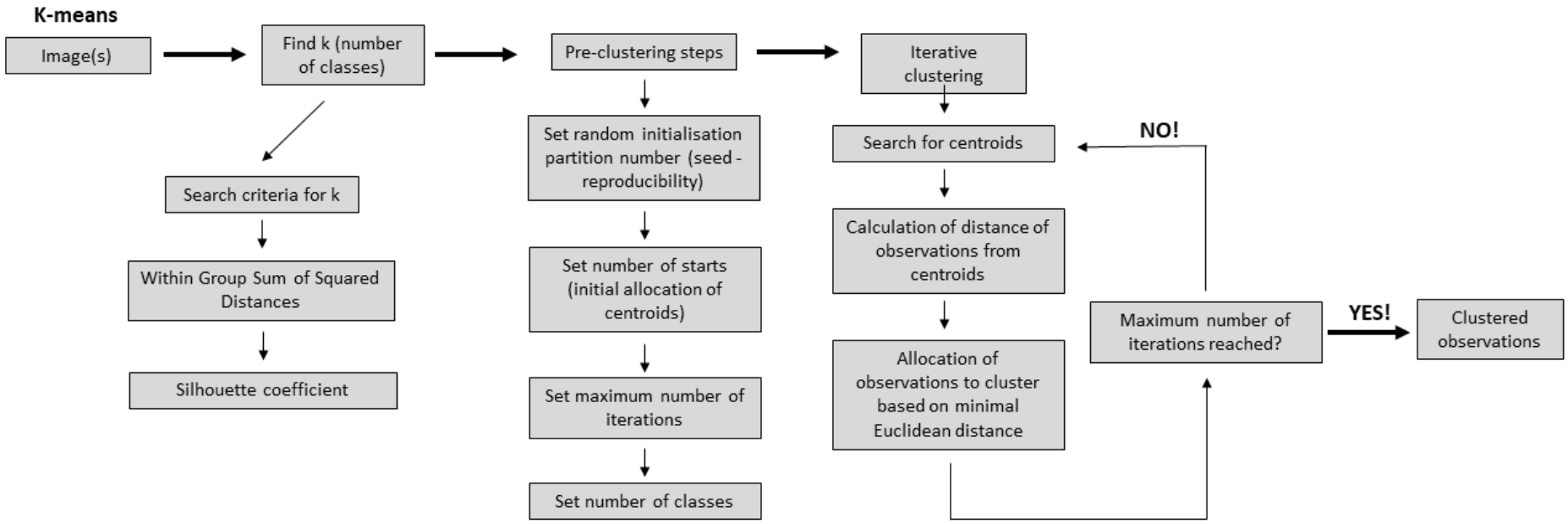


Figure 2.13 – Schematic representation of the steps involved in k-means unsupervised clustering procedure

2.4.2 Supervised image-analysis

Supervised image classification is a technique rooted in machine-learning: approaches where algorithms are trained to “learn” patterns in the underlying data and recognise those patterns in unseen data (Mohri et al., 2012). In supervised image classification ground-truth data (the response variable) are required for the algorithm to learn the underlying data structure and apply decision-rules that classify the unseen data. The ground-truth dataset is commonly divided into training and validation subsets according to various splitting and stratification (Millard and Richardson, 2015). Given that these algorithms “learn” from the data, it is important that training samples are representative of each category targeted in the prediction. It is thus good practice to assess their representativeness. First, by comparing the cumulative frequency distribution of the explanatory variables (e.g. backscatter and bathymetry) extracted at the sample locations (possibly within a buffer to account for positional errors and improving the estimation of average values) against that of the whole explanatory variable. Secondly, using box-plot analyses the class separation potential can be scrutinised (Chapters 3 and 5). Typically, the steps involved in a supervised image classification routine are as follows: 1) preparation and exploration of the sample sets, 2) computation of the information system, 3) reduction of the information system via feature selection procedures to avoid redundancy and build simpler and less computationally expensive models, 4) tuning algorithm-specific parameters (e.g. number of trees grown in a Random Forest model), 5) running the model, and 6) compute the accuracy of the prediction in respect to the withheld validation samples.

Various supervised algorithms exist and have been tested in the marine mapping literature over the past few years by a limited amount of comparative studies (Ierodiaconou et al., 2011; Diesing et al., 2014; Stephens and Diesing, 2014; Montereale Gavazzi et al., 2016). However, the most recent research focusing on supervised image classification methods pointed at the considerable performance of tree-based classifiers, of which in particular Classification and Regression Decision Trees (CART) and Random Forest (RF - Breiman, 2001) classifiers (Rattray et al., 2013; Stephens and Diesing, 2014; Diesing et al., 2014; Montereale-Gavazzi et al. 2017; Ierodiaconou et al., 2018; Misiuk et al., 2018; Porskamp et al., 2018; Turner et al., 2018) (Chapters 3 and 5). A schematic overview of the supervised routine set-up using the RF approach is given in Figure 2.14 and is further described in Chapters 3 and 5.

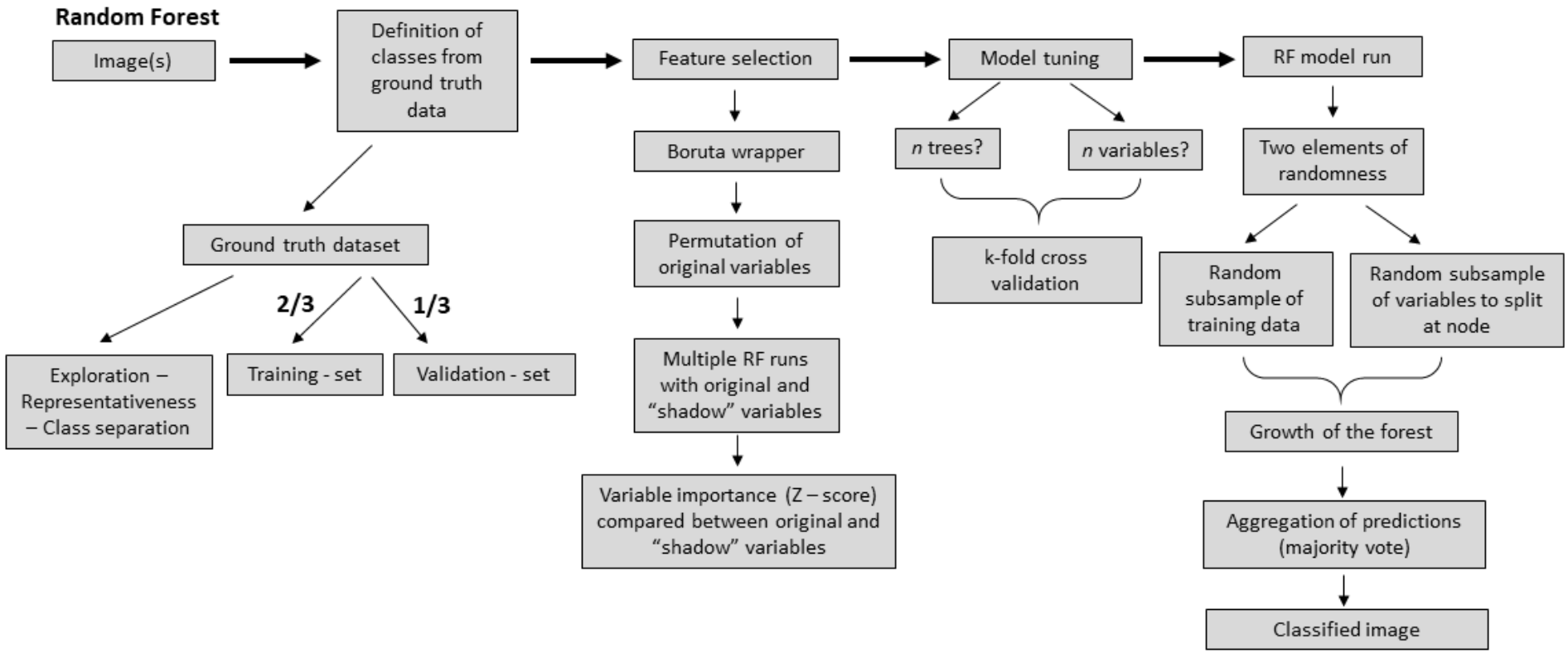


Figure 2.14 – Schematic representation of the steps involved in supervised image classification using the Random Forest algorithm. Under the "Model tuning" tab, *n trees* and *n variables* (or *mtry*), refer to the number of trees grown in the forest and the number of variables used at the split of each node. The random subsampling procedure is that of bootstrap sampling (with replacement). Combination of bootstrap sampling at rows and columns with aggregation of final votes (when RF is used for classification) or values (when used for regression) is referred to as *Bagging* (Breiman, 2001).

2.4.3 Accuracy assessments of predictive models by image-analyses

The accuracy assessment is the final and fundamental stage of an image classification study (Congalton, 1991; Foody, 2002; Pontius et al., 2004). At this stage, accuracy metrics expressing the confidence of the mapping product are derived from the confusion matrix (also referred to transition matrix when use for change detection between two images – Chapters 5). A confusion matrix cross-tabulates predicted and observed instances (Table 2.2): the thematic/categorical values of the validation samples withheld form the overall sample set are compared to the predicted values at their same location on the map. The matrix then displays observed instances over the rows and predicted instances over the columns. Initially, the $n \times n$ matrix allows to observe the agreement and "confusion" between categories and classified maps. The diagonal entries display the agreement between maps, hence the correctly classified pixels for a given category. The off-diagonal entries display the confusion between classes, for example, how many instances of category A have been predicted as category B, and so forth. Accuracy metrics relate to: (1) Overall Accuracy (A), (2) kappa statistic (Cohen, 1960) (k), (3) No-Information Rate (NIR) and (4) User (A_U) and (5) Producer (A_P) accuracies of individual classes (Congalton, 1991). Table 2.2 reports these accuracies. (1) Measures the overall agreement by considering the sum of diagonal entries divided by the overall sum of instances, (2) also considers the off-diagonal entries and measures the *"proportion of agreement after chance agreement have been removed from considerations"* (Cohen, 1960). (3) Measures the largest class percentage in the data and communicates whether the model did better than only predict the most frequent class (i.e. an overall accuracy inflated by the class that occupies more space in the dataset). (4) Measures the reliability of the prediction of a given category considering the correctly classified instances of that classes and those predicted as such whereas (5) measures how accurately a given category has been classified in respect to the overall number of validation samples for that category (from Banko, 1998 and Congalton, 1991).

Table 2.2 – The confusion matrix cross-tabulating observed (rows) and predicted (columns) instances. Italicised values along the diagonal in the 4 x 4 classes sample matrix display agreement between observed and predicted instances. Off-diagonal entries display the confusion between classes.

		Predicted					
		Category 1	Category 2	Category 3	Category 4	Σ (Rows)	User A
Observed	Category 1	*P11*	P12	P13	P14	P_{1+}	P_{1+} / P11
	Category 2	P21	*P22*	P23	P24	P_{2+}	P_{2+} / P22
	Category 3	P31	P32	*P33*	P34	P_{3+}	P_{3+} / P33
	Category 4	P41	P42	P43	*P44*	P_{4+}	P_{4+} / P44
	Σ (Columns)	P_{+1}	P_{+2}	P_{+3}	P_{+4}	N (Total)	
	Producer A	P11 / P_{+1}	P22 / P_{+2}	P33 / P_{+3}	P44 / P_{+4}		
	Overall A (P_0 – observed agreement between classifiers)	$\Sigma P_{ii} / N$					
	Chance A (P_E – expected probability of agreement)	$\Sigma P_{i+} P_{+i} / N^2$					
	kappa (Cohen's)	$P_0 – P_E / 1 - P_E$					

2.4.4 Recent investigations on seafloor mapping using automated image-analysis

Hereafter, a selection of studies presenting MBES and ground-truth data integration through automated and semi-automated supervised and unsupervised routines are presented.

One of the first published supervised image-classification benchmark comparative studies is the one of Ierodiaconou et al. (2011) where the predictive accuracy (or model performance) of Maximum Likelihood Classification (MLC - parametric) and of two new generation Decision Tree (DT) classifiers (non-parametric) were tested, integrating videographic observations and gridded MBES data and morphometric derivatives for the prediction of six habitats. They reported the considerably poor accuracy of the MLC approach (compared to the DT methods) justified by the inherent assumption of normally distributed data of this method which resulted in having limited applicability for the prediction of heterogeneous classes. Stephens and Diesing (2014) compared six supervised classifiers (Classification Trees, Support Vector Machines, k-Nearest neighbour, Neural Networks, Random Forest and Naïve Bayes), equally based on gridded MBES data and bathymetric derivatives, and used a large legacy dataset of sediment samples acquired by grabs. They identified Tree-based classifiers as outperforming the rest of the selected algorithms and reported on the importance of undertaking both feature selection routines and tuning of model parameters for the construction of simpler and more objective models. Unsurprisingly, following the outcomes of these benchmarking studies, tree-based classifiers have been reported as highly performant by, for example, Porskamp et al. (2018). In the latter study, Random Forest proved being an optimal classification routine learning complex patterns in the data structure from an input of gridded MBES layers, as well as modelled hydrological variables. In agreement to Montereale-Gavazzi et al. (2016) clear trade-offs between number of classes, number of ground-truth samples and model accuracy were identified. Similarly, Turner et al. (2018) integrated MBES gridded data and videographic ground-truth observations and reported significantly higher predictive accuracies obtained by Random Forest modelling, compared to single classification trees and Naïve Bayes classifiers.

Regarding unsupervised approaches, Eleftherakis (2013) and Snellen et al. (2018 - and references of previous work therein) used Principal Component Analysis (PCA) of backscatter statistics combined with k-means clustering, as well as applied a classification approach referred to as the Bayesian seafloor classification technique (Simons and Snellen, 2009). They equally used sediment grab samples to interpret the classified outputs. In the first approach, their classification is based on backscatter-derived statistics such as the mean, mode, skewness, minimum, maximum and standard deviation that are “condensed” (i.e. orthogonally transformed from possibly correlated to linearly uncorrelated Principal Components that maintain the variance explained) using PCA, and the set of output Principal Components explaining most of the variance used for clustering by the k-means. An important outcome of this study

is the reference made to the challenging task of statistically defining the optimal number of clusters in backscatter data, hindered by a lack of clear separation of peaks in the data-structure. The Bayesian technique has been found to produce comparable results to the previous approach. Lamarche et al. (2011) presented a study where the GSAB modelling approach was applied to a regional and calibrated backscatter dataset from which the main sedimentologic classes (mud, sand, and gravel) were directly inferred from the backscatter data: a useful approach where a paucity or complete lack of ground-truth data existed, producing general, *sensu lato* description of the sediment type, satisfying a broad scale regional assessment. However, this study benefited from the absolute calibration of the backscatter data, which could be compared to ground-truthed and calibrated measurements acquired elsewhere to understand the seafloor nature in more detail. Simple clustering of backscatter data alone has also been reported to yield highly predictive models of benthic substrates and habitat (e.g. Fogarin et al., 2019; Hasan et al., 2016; Montereale-Gavazzi et al., 2016). However, it is generally recognised that the discriminatory performance of a classification increases where bathymetry and backscatter (and derivatives) are used in combination (Eleftherakis, 2013).

2.5 Ground-truth data acquisition and processing

Hydroacoustic data provide only indirect observations (i.e. remotely sensed) of the seafloor and its nature under the form of “acoustic diversity”. As a result, it is critical that ground-truth data complement the remotely sensed data to establish links with the real world. While to a degree (i.e*. sensu lato*), physical attributes of the seafloor can be inferred from the AR (see *Empirical modelling of angular response* – Chapter 3 and 4), ground-truth data are important irrespective of the backscatter data type used. Depending on the survey area, the seafloor can be highly dynamic at various spatio-temporal scales (Anderson et al., 2008), therefore it is important that ground-truth data are collected in complement to the hydroacoustic surveys (i.e. as close as possible in time), as otherwise this may lead to erroneous interpretations of the reflectivity patterns, and erroneous interpretation of accuracy metrics may arise when, for example, statistically validating model predictions using legacy datasets (e.g. Stephens and Diesing, 2014).

There exists a multitude of sampling methods ranging from physical (i.e. grabs and corers) to optical (i.e. videographic instrumentation) gears (Figure 2.15). The use of different sampling gears will depend on the kind of substrate being sampled for and on the kind of information targeted: Van Veen grabs and box corers are generally adequate for soft and unconsolidated sediments, whereas Hamon grab and videography are best suited to harder substrates where gear-penetration into the sediment is minimal and/or not feasible (Blomqvist, 1991). Physical sampling gears concern the sampling of relatively small portions of the seafloor (i.e. local information), penetrating about 20 to 30 cm and sampling areas of about 0.1 to 0.4 m^2. Videographic methods such as the Sediment Profile Imager (SPI) and camera systems installed on drop-frames and or towed-frames can be employed to acquire images of the water-sediment interface (e.g. Assis et al., 2007; Montereale Gavazzi et al., 2016). Additionally, Box-corers and SPI samples can provide useful information of shallow subsurface layering/configuration.

The sampled footprint has implications for the integration with hydroacoustic data using image analysis given that, ideally, the support size should match in both data types (sampled surface area *versus* pixel size). In this regard, collection of replicate samples (Chapter 5) can aid in improving the inter-agreement between the data types support sizes, while also ensuring the substrate homogeneity of the acoustic facies being sampled. However, it must be noted that ground truthing is highly labour-intensive and time-consuming, often leaving mappers the choice of either collecting fewer samples with replicates or more samples without: choices must be made based on ship-time availability and on the hydro-meteorological conditions, both expected for the duration of the campaign and at the time of sampling. Depending on the water clarity/turbidity, dictating the height above seafloor at which video systems can be operated to obtain good quality images, videographic sampling approaches

can better approach the scale of the acoustic imagery (e.g. Gazis et al., 2018 - Chapter 3 and 5) and are considerably less invasive.

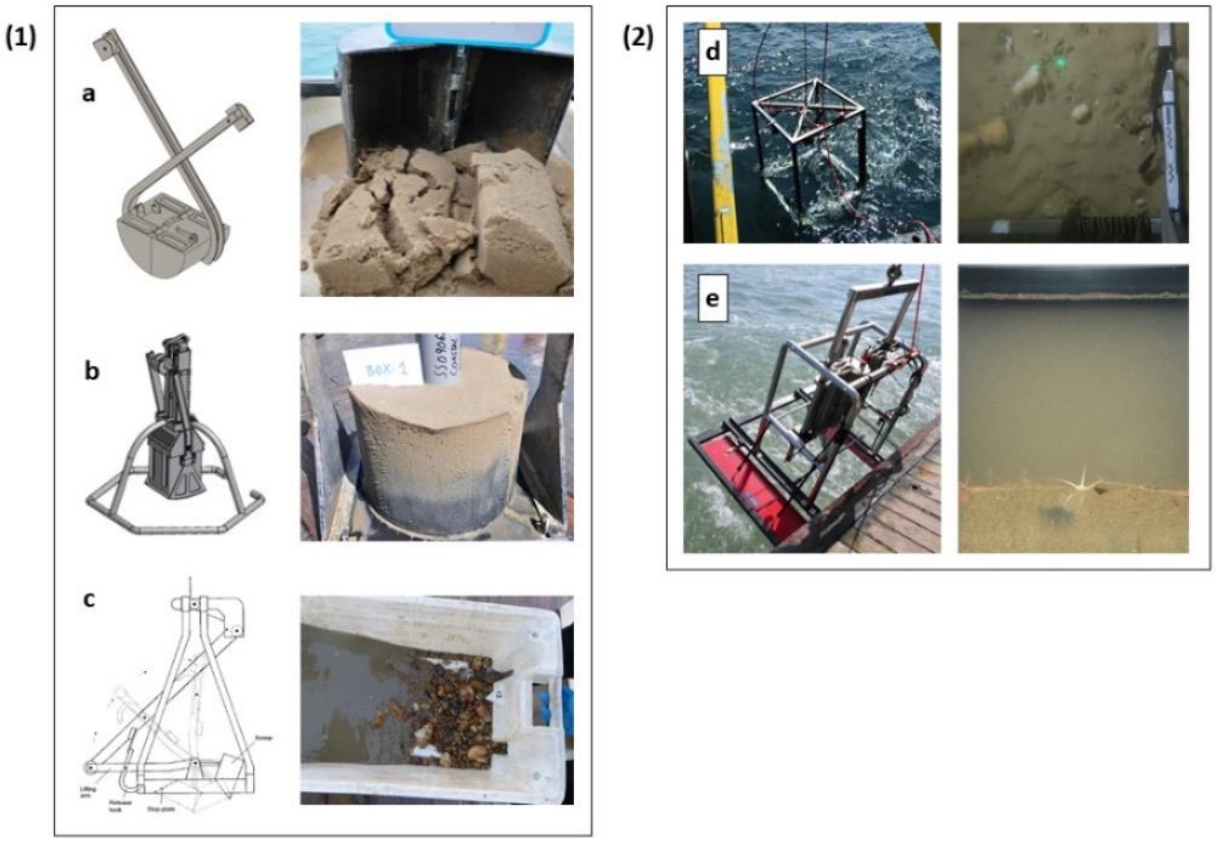


Figure 2.15 – Summary of the ground-truth gears operated on board *RV Belgica* and *RV Simon Stevin* throughout this thesis work. Panel (1) displays the physical sampling gears: a) Van Veen grab, b) Box-corer and c) Hamon grab. Panel (2) displays the videographic sampling gears: d) Video drop-frame and e) Sediment Profile Imager. Pictures of the type of sample acquired by each gear are displayed aside each instrument.

The sediment bulk retrieved by physical sampling gears is generally analysed in terms of granulometry by means of laser diffraction for the fine sediment fraction (grains diameter < 1 mm), and by sieving through different sized meshes for the entire or coarser sediment fractions (> 1 mm) which are in turn individually weighted. Weight percentages and statistics of the sediment distributions (such as the median diameter – D50 and the percentage of given sediment fractions) are derived as means to seek quantitative relationships with the hydroacoustic data (Chapters 3 and 4). During the collection of samples, an on-board protocol for describing the sediment attributes visible at the sample's surface is used by trained geologists and/or sedimentologists (Figure 2.16 gives an example of the one used throughout this work). While qualitative, this information can replace the lack of data when an insufficient amount of sediment bulk is collected for sieving. Importantly, the qualitative information of the coarser sediment fraction (i.e. type and %) can be obtained, having important effects on the backscatter response, hence critical for any backscatter classification study.

The results of the sedimentological analyses are commonly summarised into categorical groups according to existing sediment classification schemes such as the Wentworth (1922) and Folk (1954) schemes, capturing the full spectrum of sediment types, from fine sediments such as clay and mud, to coarser sediments such as sand and gravel.

Name: .....
Institute: .....
Project: .....
Date: ..... **Time:** .....

museum

Nr. .....

Gear: Grab Sample Box Core
Vibro Core SPI
Other: .....

Lat/y: ....
Lon/x: ....
Depth/z: ....

Ship: .....

Picture: ☐ **Grain size analyses:** ☐

SEDIMENTOLOGICAL ANALYSIS

MAIN COMPONENT ..... %

..... - ..... m	clay (mud)	silt (mud)	fine sand	medium sand	coarse sand	gravel	peat
..... - ..... m	clay (mud)	silt (mud)	fine sand	medium sand	coarse sand	gravel	peat
..... - ..... m	clay (mud)	silt (mud)	fine sand	medium sand	coarse sand	gravel	peat
..... - ..... m	clay (mud)	silt (mud)	fine sand	medium sand	coarse sand	gravel	peat
..... - ..... m	clay (mud)	silt (mud)	fine sand	medium sand	coarse sand	gravel	peat

Colour: .....[text] .....[text]

Remarks: [grain-size, inclusions, lenses, layers, laminae, coarsening or fining upwards, heterogeneity, mixing sediment, etc.]
.....
.....
.....
.....
.....

SECONDARY COMPONENT ..... %

	clay (mud)	silt (mud)	sand	gravel	OM/peat	shells	glauconite
High	☐	☐	☐	☐	☐	☐	☐
Normal	☐	☐	☐	☐	☐	☐	☐
Low	☐	☐	☐	☐	☐	☐	☐

Remarks: [amount, size, composition, inclusions, lenses, layers, laminea, black spots, etc.]

BIOLOGICAL ANALYSIS

SHELL COMPONENT ..... %

SHELLS: grit fragments shells .....

Species: .....
.....
.....
.....

Remarks: [amount, size, composition, bioturbation, reworking, etc.]
.....
.....
.....

CHEMICAL ANLYSIS

A-RPD [APPARENT REDOX DISCONTINUITY LAYER]

Depth: .....[MIN] .....[MAX] *continu / discontinu*

Remarks: [anoxic black layers, anoxic black spots, mixing sediment, reworking, bioturbation, added material, etc.]
.....
.....
.....

OTHER REMARKS
.....
.....
.....

SKETCH

Figure 2.16 – Template example of the descriptive protocol used during the sampling campaigns on board RV Belgica and RV Simon Stevin throughout the timespan of this doctoral work.

Videographic samples, from which still frames are typically extracted, require annotation of the images into descriptive schemes (e.g. Rooper and Zimmermann, 2007; Yoklavich et al., 2000). Annotation is achieved either by expert interpretation (grid count) or by automated image-analysis methods (pattern recognition). Underwater laser pointers or rulers are needed to provide an understanding of the sizes of the objects observed within the images, and for the subsequent conversion of pixels into metric units. A severe limitation of drop-frame videographic samples, at least for areas with significant hydrodynamic forcing (such as the BPNS), is the fact that the acquisition time is constrained by the slack-water window (~ 1 h) when the SPM transport in the water is reduced and the visibility improves. This also has implication when applying automated image-analysis approaches in the derivation of semi-quantitative descriptors such as percent gravel, bare ground and biota (Figure 2.17 – Chapter 3).

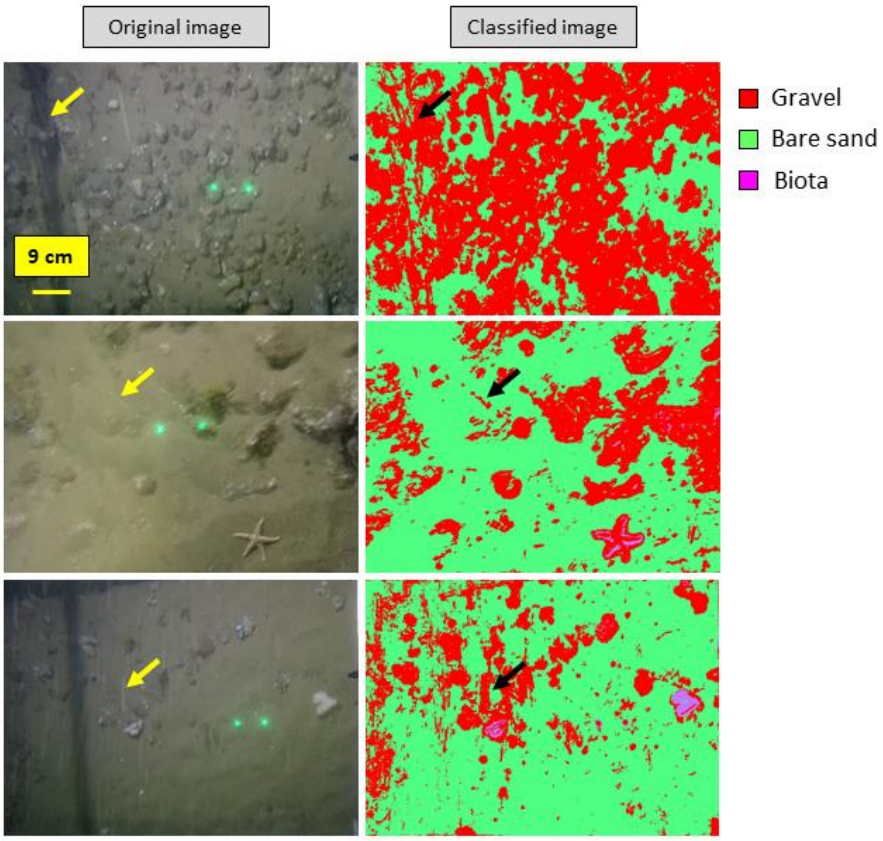


Figure 2.17 – Example of automated extraction of substrate descriptors (trials) using ImageJ (https://imagej.nih.gov/ij/) software built-in image processing tools. Left and right panels display original and classified images respectively using a two classes scheme for gravel and bare ground. Thanks to the metric reference scale given by the laser pointers, the size of the frame and therein contained objects can be estimated and the percentage cover of gravel and/or of bare sand, converted to metric units. Image-classification allows to derive percentages of, in this case, gravel cover. Automated extraction of biotic cover is also possible using this approach. Yellow and black arrows evidence some artefacts resulting in the classification: in the top row, the frames' shadow results in the classification as gravel percentage. In the second row, the laser beams illuminates suspended particulate matter, promoting backscattering and resulting in misclassification and in the third row, illuminated suspended particulate matter also affects the classification process.

2.6 Backscatter for discovery, backscatter for monitoring

Acquisition of MBES backscatter can be categorised into two main types depending on the scope of the survey: (1) “mapping for discovery” (i.e. snapshot in time mapping), based on a single pass over a given study site, and (2) “mapping for monitoring” (i.e. multi-pass surveys), consisting of repeated/serial surveys of the study area (Lucieer et al., 2018). Different technical constraints are presented for these two kinds of applications. These relate to measurement calibration, accuracy, stability and repeatability (Lurton and Lamarche, 2015 – see *Backscatter calibration, repeatability and standards in acquisition and processing*).

2.6.1 Mapping for discovery

In mapping for discovery, the primary aim of the survey is exploration of a seafloor area, often unknown, and collection of hydroacoustic and ground-truth data to characterise it in terms of its geology, sedimentology and biology. A large variety of national seafloor mapping programmes have been initiated (see Chapter 3 where an exhaustive list is presented) in response to the realisation that only about 12 % of the seafloor has been spatially explicitly mapped, most of which in the coastal seas (i.e. at resolutions adequate for ecological management). Stating Diaz et al. (2004) (Chapter 3), the basic premise towards a more responsible stewardship of the marine environment relates to the following question: *“How can one accurately evaluate the relative value, in a temporal and/or spatial sense, of a specific habitat when no attempt has been made to objectively define the type and extent of the habitat itself?”*. Therefore, exploration mapping is in this case not strictly bounded to technical constraints such as backscatter data calibration, accuracy and stability (although in the long-run, when the need to merge disparate datasets to produce large geographical coverage and/or change detection, all datasets would enormously benefit from such a targeted standardisation). The main target is the characterisation of an area for the first time and the primary concern of the surveyor and user is the production of a good quality backscatter mosaic for classification.

2.6.2 Mapping for monitoring: the fourth dimension

Mapping for monitoring implies capturing the variability of the seafloor as a function of the fourth dimension: time (i.e. *x, y, z* or dB and *t*). It entails the repetition of MBES surveys over the same area of interest with the aim of monitoring the morphological and sedimentary evolution via the analysis of bathymetric and backscatter data (Fig. 2.18). Change detection, i.e., *“the processes of identifying differences in the state of an object or phenomenon by observing it at different times”* (Singh, 1989), is strictly bounded to technical constraints of backscatter measurements’ accuracy, stability and calibration as well as knowledge of the system and of its natural dynamics: together dictating their repeatability and hence the confidence associated to the detected changes based on the direct use of the backscatter measurements (where changes in dB values are directly used as a proxy for change – this would be equivalent to the

terrestrial remote sensing realm when calibrated serial measurements are relied upon; e.g. Singh, 1989).

As previously iterated, absolute measurements are not strictly mandatory as long as the system's measurements are stable and accurate for the duration of the monitoring programme (Roche et al., 2018). This is because the basic premise in directly using MBES backscatter data for detecting changes in seafloor composition is that *"changes in substrate cover must result in changes in backscatter values and changes in backscatter due to seafloor cover change must be large with respect to changes caused by other factors"* (adapted from Singh, 1989). It is exactly those "other factors", that the maritime remote sensing community needs to address to improve, or rather begin, to fully exploit the currently available data, and the undoubtably growing volumes of datasets that will become available in the near-future.

The seafloor is in a state of permanent flux at multiple spatiotemporal scales and our knowledge of this dynamicity remains scarce. While approaches to mapping (i.e. single pass survey and classification) have considerably evolved over the past decade, introducing an array of classifiers and comparative studies, as well as advancing the number of features that included in a classification problem can enhance discrimination and the detail achievable by the classification, serial surveys and change detection remains a premature application, requiring investigators to test and asses change detection methodologies for which there is a paucity of studies (e.g. van Rein et al. 2011; Rattray et al., 2013) and importantly, begin the build-up of baseline knowledge pertaining the natural variability at different spatio-temporal scales (e.g. Ernsten et al., 2006); this refers to both the intrinsic type of variability, associated to the seafloor itself, and to the unwanted, "exogenous" variability; the kind of variability that influences the acoustical measurements and can hinder the direct use of the remotely sensed data. Because of these reasons, and the challenging operational environment, underwater change detection should be seen as amongst the greatest challenges in the acoustic seafloor classification discipline, making it indeed, a very exciting topic with several years of research to come.

To begin these efforts, an array of instrumental and environmental sources of variation must be kept in mind at various stages of the acquisition and the processing chain of the backscatter measurements (Lurton and Lamarche, 2015; Lucieer et al. 2018). During acquisition, important factors to be considered are:

1. Sea-state linking to the carrier platform motion, for example creating under-keel bubbles when squatting, and linked to the IMU compensation (pitch and roll beam stabilisation). Addressed by conducting surveys only under favourable meteorological conditions.

2. Unchanging the acquisition parameters (e.g. frequency, pulse length, beamforming mode). Addressed by maintaining rigorous operational standards.

3. The navigation heading in respect to the underlying seafloor morphology (i.e. azimuth dependence) and the polarisation of small-scale roughness (beyond the imaging-capability of the sensor) driven by the local hydrodynamic regime. Addressed by carefully planning survey strategies and maintaining rigorous navigational standards.

4. The seawater hydrological status. Addressed by the characterisation of the local hydrological conditions and computation of correction empirical coefficients.

5. Temporal drift of the system: i.e. aging of the antennas and biofouling. Addressed by regular dry-dock and/or diver-based maintenance operations.

During post-processing, it is of uttermost importance to use one processing software only, in order to maintain a consistent processing workflow for the entire monitoring dataset.
Figure 2.19 shows a summary of the confounding factors that must be considered when acquiring backscatter and addressing change detection in the marine environment.

Furthermore, it is important to control the measurement stability and repeatability (perturbed by instrumental and environmental drift), whether or not working with relative and/or absolute calibration of the sensor (Lurton and Lamarche, 2015 – Table 2.3). As previously iterated, an operationally viable and increasingly accepted emerging solution is that of controlling the measurements over a well-known and stable seafloor reference area (for the BPNS, see Roche et al., 2018 for a detailed account). From an operational perspective, this is a pragmatic solution for hull-mounted (i.e. fixed) uncalibrated echosounders (Eleftherakis et al., 2018). The advantages of surveying such reference areas at the start of oceanographic missions are manifold and include the possibility to compare measurements from different echosounders and platforms (operating at a single frequency) in space and time (Hughes-Clarke et al., 2008; Roche et al., 2018; Weber et al., 2018) (Chapter 3) and allow an efficient control of the stability, hence the repeatability of the backscatter measurements of a given system to detect changes (Table 2.3).

Table 2.3 – Definitions of accuracy, stability and repeatability of backscatter measurements: addressed using a natural, well-documented and stable reference area (for Accuracy and repeatability). Stability is addressed by careful control of environmental and instrumental sources of variance at the acquisition and processing phases of the backscatter measurements (from: Lurton and Lamarche, 2015; Roche et al. 2015 and Eleftherakis et al. 2018).

Parameter	Definition
Accuracy	Estimation of the measured level of uncertainty with which to provide the ability to detect changes in the seafloor environment over that of mapping uncertainty
Stability	Monitoring and control of variability and sources causing discrepancy between serial measurements
Repeatability	Quantitative comparison between serial surveys over for example a reference area (comparable to a bathymetric patch test)
Absolute calibration	Reduction of dB values to absolute calibrated measures, reflecting intrinsic and univocal physical characteristics of a given target (i.e. metal sphere targets as in calibration methods of the fisheries acoustics field)
Relative calibration	Relative dB values, yet consistent in respect to measures acquired over a reference target of know nature - comparable measurements of one system operated at different times over *n* surveys
Calibration propagation	Whether absolute or relative, the propagation of the calibration compensations from echosounder to echosounder, initiated by considering a nominal truth from which to obtain the calibration bias

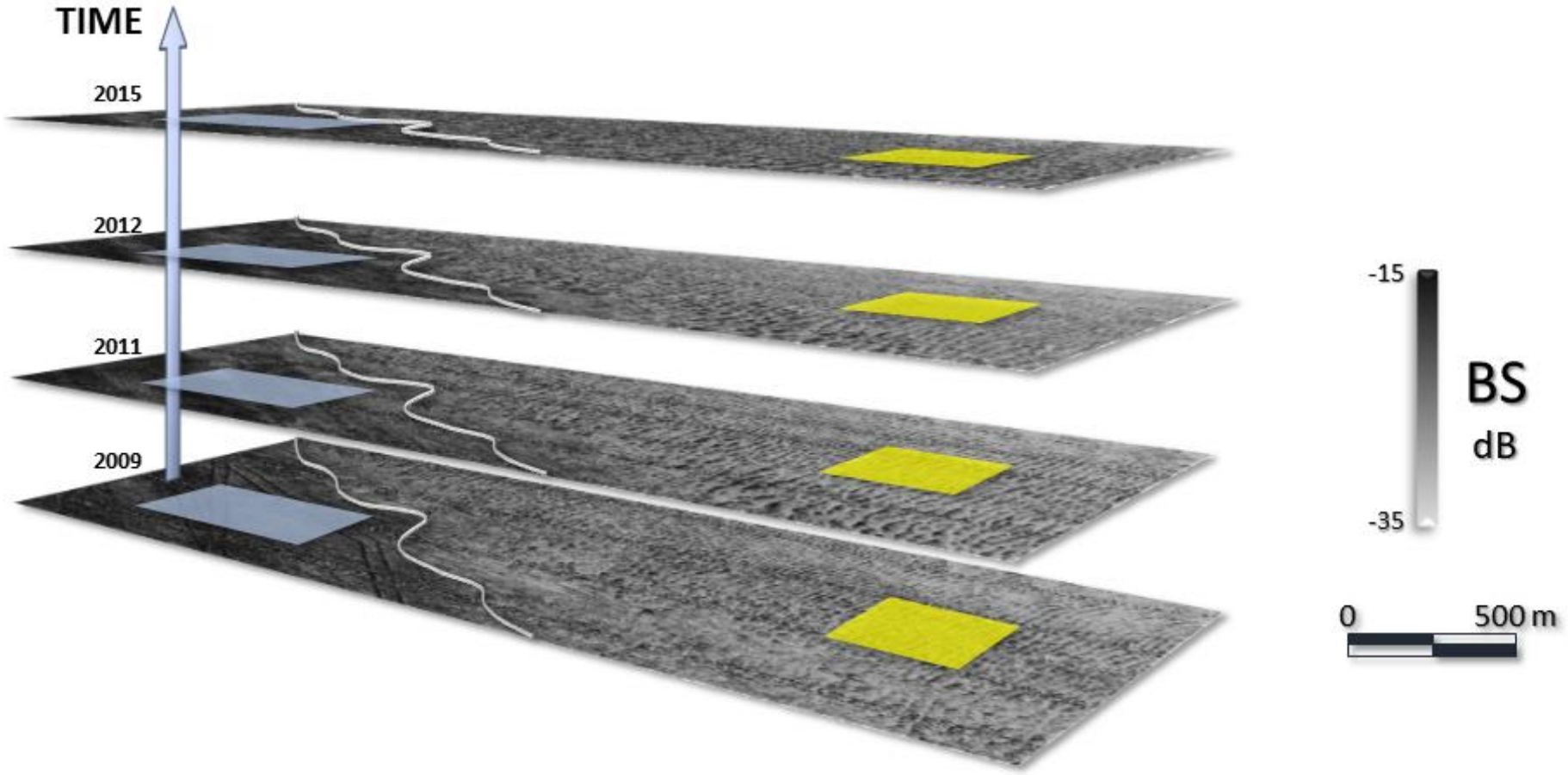


Figure 2.18 – Schematic representation of a multi-pass MBES backscatter time-series survey dedicated to monitoring of trends, patterns and physical changes of a seafloor area. (Image readapted from https://economie.fgov.be/sites/default/files/Files/Entreprises/Sand/13-GEOHAB2015-Presentation.pdf).

Minimally, change detection can be subdivide into pre- and post-classification. In the former case, serial datasets must be radiometrically corrected so as to avoid the detection of changes that are not due to seafloor changes (e.g. hydrological status during the acquisition of a survey; comparable to the reflectance disparities arising from satellite images acquired during different sun illuminating angles or during different atmospheric conditions – Coppin et al., 2004). In post-classification, changes are assessed by comparing thematic models, thus deriving from-to transitions between categorical themes within two or more scenes. An approach does not exclude the other, and the paucity of studies of this topic in the maritime remote sensing community, leaves the door open to experimentation and comparison of methodologies where cross-evaluation of approaches is largely needed (Chapter 5). As previously discussed, besides the use of CBI, AR backscatter can also represent a form of change detection information, as much as it represents a critical classification feature in the task of acoustic classification in its own respect: clearly, we are at the very nascence of these approaches and there is a great need to test, explore and cross-evaluate both the methodologies and the dynamicity of the environment we ultimately target to monitor.

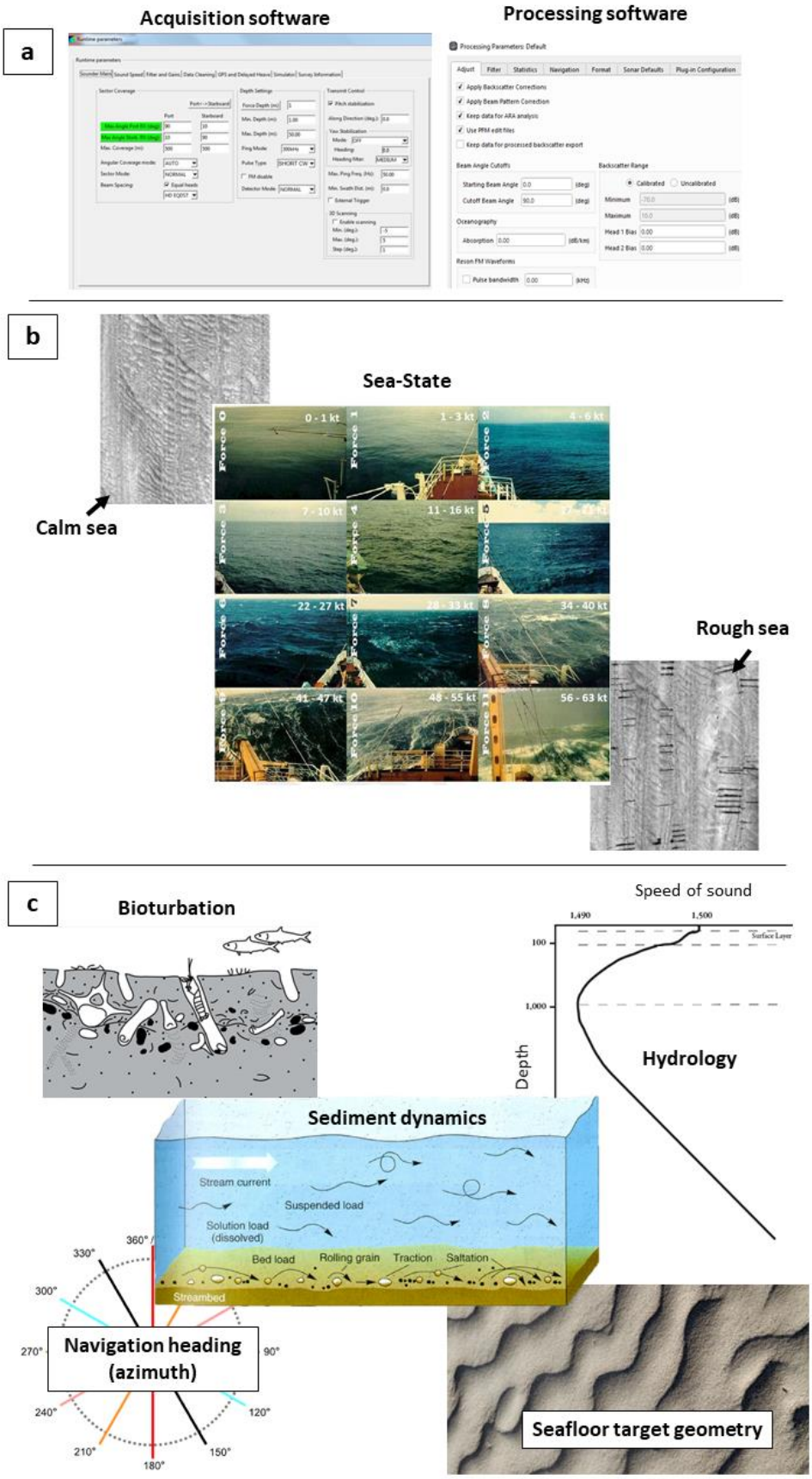


Figure 2.19 – The three main categories of sources of variability in MBES seafloor backscatter. Caption continues on next page.

a) Acquisition and processing software parameters must be rigorously standard within a monitoring program/change-detection study. b) The sea state affects the motion compensation and severe vessel motion such as pitch can lead to creation of under-keel bubbles, resulting in a significantly reduced SNR. Note the gridded backscatter acquired in “rough sea” conditions and presenting “stripy” artefacts. c) Environmental sources of variance can be related to the effect of bioturbation, hydrological conditions including the effect of turbidity, the seafloor target geometry may change in orientation and/or shape under the effect of the local hydrodynamic regime and lastly, the navigation heading in respect to the underlying morphology (in particular the small scale structures beyond the sonar imaging capability) can severely affect the backscatter response resulting in an azimuthal dependence. (adapted from: b) https://blog.metservice.com/wp-content/uploads/2015/03/Fig-4.jpg and Lurton and Lamarche, 2015 ; c) centre: http://geog.uoregon.edu/shinker/geog101/lectures/lec08/lec08_figs/fig14-12.gif; clockwise starting in the top left corner Boessenecker et al., (2014); https://dosits.org/wp-content/uploads/2018/01/profile2-build2018-600x778.png;
https://freerangestock.com/sample/37502/desert-sand-ripples-texture.jpg; Lurton et al., (2018).

Chapter 3

3. Integrating multi-source multibeam and ground-truth data to seamlessly map continental shelf substrate types.

Application to the Belgian Part of the North Sea

3.1 Abstract

Detailed and accurate information of the spatial and temporal distribution of seafloor sediments is of great relevance for several marine management and industrial applications. Multibeam echosounder systems (MBES), acquiring co-located bathymetry and backscatter data, have become the mainstream tool in seafloor mapping, and drastically improved the spatial resolution of traditional seafloor maps. With increasing volumes of data becoming available and the need to inform environmental policy-making with environmental spatial datasets, the need comes to develop strategies to produce sediment maps via objective, repeatable and accurate classification methods. In this study, ~150 km^2 of 300 kHz multi-source MBES surveys and a set of 163 ground-truth samples, acquired across the Belgian Part of the North Sea over the period 2015-2018, were integrated to predict seafloor sediment type. Based on survey overlap on a natural reference area, a strategy to merge the disparate data into a seamless gridded reflectivity and bathymetry map was implemented. This dataset was used to explore relationships with sediment type and to predict its distribution across the entire survey extent. Routines based on unsupervised k-means and supervised Random Forest classifiers were tested against their predictive accuracy, including the effect of prescribing two commonly used classification schemes at the European level (European Nature Information System level III and Folk sediment categories). Results indicate that the modulation of the average seafloor backscatter intensity relates to the relative proportion of the sediment fractions; linearly decreasing with increasing percentage of the fine- to medium sand fraction (0.063 - 0.5 mm grains) and increasing with increasing percentage of the coarse sediment fraction (0.5 - > 10 mm). Furthermore, within well-sorted and homogenous fine sand (0.063 - 0.5 mm) seafloor areas, backscatter intensity is positively correlated to the median diameter of the grains. However, for poorly sorted and rather heterogeneous areas, presenting mixtures of coarse sediment fractions, ambiguous relationships are found. No 1:1 relationship can be established between acoustic backscatter and Folk class sediment type: pointing at a lesser discriminative ability for coarser sediments. Class aggregation is needed to fit the prescribed classification scheme to the remotely-sensed data. Under these constraints, unsupervised clustering of backscatter data was found to poorly perform for both classification schemes, particularly with increasing number of classes. Supervised Random Forest produced highly accurate results at both classification levels. Overall the thematic map accuracy ranged from 0.44 to 0.85 and chance agreement (Cohen's k) from 0.28 to 0.76. In line with recent literature on predictive seafloor mapping, Random Forest classification is confirmed as an optimal approach providing satisfactory, objective and repeatable results. Analysis of the angular response backscatter at the sample locations further comforted the RF classification showing overall good class distinction between the mean curves for each sediment category, as well as confirming their physically meaningful differences. The maps produced are essential in advancing the spatially-explicit understanding of the

seafloor and can be used to guide benthic habitat mapping studies of the largely unexplored gravel beds of this North Sea region.

Keywords: Multibeam backscatter, substrate mapping, classification, multisource surveys, ground-truthing.

3.2 Introduction

Coastal and marine ecosystems are regarded as amongst the most productive and valuable environments on Earth (Barbier et al., 2011; Guelorget and Perthuisot, 1992). Overall, coastal regions are subject to an increasing occupancy by human populations and thus, are highly vulnerable to a multitude of increasing anthropogenic pressures (Halpern et al., 2008). While ~ $ 1.5 trillion of the yearly total global economy are generated by commercial exploitation of "ocean products" (OECD, 2019), merely 5-10 % of the global seafloor has been mapped at resolutions adequate for ecological management (Sandwell et al., 2006). In response to the global realisation of a deteriorating health of the known marine environment, large programmes have started to map the seafloor within their Exclusive Economic Zones (EEZ), updating current sediment and habitat distribution maps through the advancing discipline of Acoustic Seabed Classification (ASC) (Anderson et al., 2008; Brown and Blondel, 2009). Various examples exist of which some significant efforts follow: Mapping European Seabed Habitats (MESH - searchmesh.net), MESH Atlantic (meshatlantic.eu), EU Seamap (jncc.defra.gov.uk/page-5040), the Norwegian MAREANO (mareano.no/en), UK Seamap (McBreen et al., 2011), the Irish INFOMAR program (infomar.ie), the Gulf of Maine Mapping Initiative (gulfofmaine.org/gommi), the Victorian marine habitat mapping project in Australia hdl.handle.net/10536/DRO/DU:30010514) and the Italian RITMARE (http://www.ritmare.it/en/) and ADRIPLAN (adriplan.eu/) initiatives.

These initiatives are strictly bounded to national and international legislation dedicated to the acknowledgment and management of marine natural resources (i.e. resource mapping and Marine Spatial Planning; Douvere et al., 2007, 2008) and the ecosystem services we depend upon (Diaz et al., 2004). Among other European legislation, the European Marine Strategy Framework Directive (MSFD, 2008/56/EC) mandates European Member States to monitor the state of their national waters and to implement management strategies dedicated to the achievement of a Good Environmental Status (GES - Rice et al., 2012). GES is addressed by eleven descriptors (including one on Seafloor Integrity – D6) and each Member State formulates assessment indicators (see Rice et al., 2012). As a result, a regional mapping is underway in Belgian waters and the Belgian State (2012) formulated two seafloor integrity-related indicators for which multibeam echo sounding (MBES) was selected as the mapping and monitoring technology.

The first step in assessing anthropogenic impacts on marine ecosystems is the acknowledgment and identification of the spatial variability and distribution of benthic habitats and/or of their abiotic surrogates (Diaz et al., 2004). There are several interpretations of the term "habitat" (Hall et al., 1997), though the term can be minimally referred to as the combination of abiotic and biotic conditions that together promote occupancy by communities of given benthic biota. Since most infaunal benthic organisms inhabit the top 10 cm of the seafloor sediment (Miller et al., 2002)

and epifaunal communities rely on the structural complexity of the water-sediment interface (Hewitt et al., 2005), detailed mapping of surficial sediment type distribution (i.e. benthic substrates) is recognised as a fundamental ecological descriptor (Kostylev et al., 2001; McArthur et al., 2010). Traditionally, sediment mapping has been achieved by means of in-situ observations (namely grab and core sampling) that are interpolated and extrapolated (Stephens and Diesing, 2014). Whilst such approaches reveal substantial and valuable spatial information (e.g. Verfaillie et al., 2006), the overall density and coverage of the sampling efforts are rarely sufficient to depict the often-complex distribution of seafloor sediments (i.e. metric-scale heterogeneity and patchiness). For example, habitat edges and the morpho-sedimentological relationship, will remain largely unaccounted.

Since the late 80's, high-frequency MBES have superseded the former mapping instruments (namely single-beam echosounders and side-scan sonars [SBES-SSS]) and have become the instrument of choice for underwater mapping because of their ability to cost-effectively co-register precisely georeferenced bathymetric and backscatter intensity data over relatively large portions of the seafloor (~ 5x water depth), and at considerably higher resolutions (~ 2% water depth at nadir) than traditional mapping approaches (depending on system configuration, in the order of centimetres for high-frequency sonars operated in very-shallow water (≤ 10 m), to tens of meters for systems operating in deep water areas; De Moustier, 1986; Kenny, 2003). The local information of the ground-truth samples can thus be extended (i.e. predicted by proxy) to continuous coverages of substrate and habitat types using remotely-sensed hydroacoustic data. The bathymetry is the primary data of a multibeam echosounder and it describes the geometry of the seafloor derived from measures of the echo's times and angles. This data type has for long received the attention of hydrographic and mapping programmes, resulting in the establishment of well-accepted international standards for acquisition, processing and accuracy estimation (i.e. International Hydrographic Organisation; Wells and Monahan, 2002). On the contrary, acquisition, processing, analysis, interpretation and quantification of uncertainties in MBES backscatter data currently stand as topical issues in the global marine mapping community which strives to reach harmonisation of approaches due to a variety of manufacturers and processing software, each implementing their proprietary adjustments to the data acquisition and processing chain (Schimel et al., 2018). As a result of these communal interests, a first set of guidelines and recommendations have recently been proposed (Lamarche and Lurton, 2018; Lucieer et al., 2018; Malik et al., 2018).

The backscatter strength (BS) reflects the amount of acoustic intensity scattered back to the sonar receiver following a complex interaction of the transmitted signal with the seafloor. It is the result of a combination of several physical factors: the water-seafloor impedance contrast (acoustic impedance is the product of density and sound velocity), the interface roughness and the sediment volume inhomogeneity, the signal incidence angle on the seafloor and the acoustical signal frequency

(Lurton, 2010). Due to the various scattering properties of different seafloor substrates, backscatter can help determine bottom type (e.g. Collier and Brown, 2005; Ferrini and Flood, 2006; Goff et al., 2004). Put simply, using MBES backscatter for sediment characterisation can be interpreted as the identification of "the characteristics and spatial organization of seafloor patches and or/signatures with common acoustic responses and the measurable characteristics of this response" (modified after Lamarche and Lurton, 2018).

MBES backscatter is generally considered at two processing levels: (1) the angular response (AR) and (2) the compensated backscatter imagery (CBI). In turn, acoustic seabed classification (ASC) can be achieved using signal-processing, via modelling of the angular response (i.e. physical process, geoacoustic and phenomenological-modelling approaches; see Lamarche et al., 2011), or via image analysis (using the CBI) which combines various supervised and/or unsupervised classification algorithms with ground-truth data (e.g. Diesing et al., 2014). Hybrid approaches combining both data types and analytical methods are also possible and are the result of very recent research (Hasan et al., 2014; Alevizos and Greinert, 2018). The AR relates to the variation of backscatter intensity with angle of incidence and is retained as an intrinsic seafloor property reflecting physical quantities of interest (Jackson et al., 1986; Lamarche et al., 2011). The variation of intensity with angle of incidence results from the dominant acoustic phenomena occurring along the angular domains of ensonification: high-intensity specular reflection around the nadir and lower-intensity scattering at oblique angles, strongly decreasing at low grazing angles. The backscatter mosaic is a derivative of the AR where BS levels are presented in a georeferenced frame, usually in the form of a gray-scale image with the angular dependence removed via statistical compensation in such a way that the whole seafloor scene seems to be observed from one same incidence angle. This is generally obtained by normalizing the data and referencing it to a conventional angle or a limited range of angles. Typically, this is in the oblique region, around 45°, where the angular dependence is weakest and where the sediment response (i.e. roughness and volume) dominates (Lamarche et al., 2011). Fundamental differences exist between these two data types in terms of type of information and spatial resolution. The angular response production requires averaging a set of consecutive pings and thus the resolution approximates the size of the swath and/or the region of interest selected for extracting the curve. On the contrary, CBI can be gridded as function of the bathymetric resolution, providing spatially explicit information of the patterns of distribution of seafloor sediments.

The dependence of MBES backscatter on the incidence angle of ensonification has led to the formulation of several signal-based approaches to seafloor classification. These are generally referred to as Angular Range Analysis (ARA) and can be subdivided into geophysical and empirical modelling approaches. A well-established modelling example using the AR backscatter is that of the composite-roughness geoacoustic model developed by the Applied Physics Laboratory of the University of

Washington (APL-UW, 1994) for low to medium-high (< 100 kHz) frequency backscatter. A range of studies have employed such a modelling approach and have demonstrated the links between sediment grain-size parameters and the backscatter AR (De and Chakraborty, 2011; Fonseca et al., 2002; Fonseca and Mayer, 2007). It must be noted however, that these models have limited applicability to a global scenario in what the naturally occurring variability and complexity of marine sediments is such that it currently cannot be encompassed into a single model for a rigorous inversion. Because of this complexity (which is a rule in marine sediments rather than an exception), the acoustic-sediment relationships have been alternatively investigated via empirical approaches. Several studies have demonstrated empirically the relationships of backscatter strength with sediment grain size (e.g: Collier and Brown, 2005; De Falco et al., 2010; Ferrini and Flood, 2006; Goff et al., 2000, 2004) and subsequently, the potential for classification (e.g: Lamarche et al., 2011; Lucieer et al., 2013). Generally, backscatter strength has been found to vary as a function of the relative proportion of fine and coarse sediment fractions, positively correlating with an increasing coarse fraction and inversely with increasing of the finer one (De Falco et al., 2010; Goff et al., 2000). Further, moderate to strong positive linear trends have also been found for the sediment median grain size (Collier and Brown, 2005; Ferrini and Flood, 2006). Recognising the complexity of these relationships, Ferrini and Flood (2006) explored the multivariate relationships between backscatter and sediment properties, including variables of microroughness, and reported strong empirical evidence of the backscatter dependency on an array of sediment and geometric variables. More recently, ambiguities in the magnitude of increase (deviation from linearity and establishment of a plateau) in backscattering strength with increasing median grain-size diameter and increasing content of the coarse fraction, have been noted in, amongst others, Snellen et al. (2018) and Gaida et al. (2018) who related the behaviour to a transition in scattering regime when the acoustic wavelength approaches and exceeds the diameter of the sediment grains (e.g. ~5 mm @ 300 kHz assuming a sound velocity of 1500 ms-1). Given the general realisation of the complexity of the acoustic response and the non-ubiquitous applicability of physical geoacoustic inversion models, empirical models have been proposed as an attempt to describe sediment type based on a restricted set of statistical parameters, relatable to physical quantities of interest. One particularly successful application is that of the Generic Seafloor Acoustic Model (GSAB). Through a set of statistical distributions, the AR can be fitted and capture the main physical processes of backscattering at incidence angles, which are in turn relatable to the underlying geological nature of the seafloor.

Classification using angular response is developing at a fast pace and recent studies have indicated the great potential for seafloor classification using such an approach (e.g. Lamarche et al., 2011; Alevizos and Greinert, 2018), in particular by providing a means to obtain general insights into the seafloor physical status in an unsupervised manner, of major interest where there is a paucity of ground-truth data (Lamarche et

al., 2011; Fezzani and Berger, 2018). However, use of the angular response remains in its infancy whereas image-based approaches are considered as better established, benefitting from the experience matured in the terrestrial remote-sensing realm. Image-based approaches in the marine mapping literature have gained popularity in the past two decades (Stephens and Diesing, 2014) and there has been a considerable effort in publications presenting data integration via semi-automated unsupervised and supervised approaches, testing pixel- and object-based image classifications and comparing various classifiers for their predictive accuracy (Diesing et al., 2014; Calvert et al., 2015, Montereale Gavazzi et al., 2016). These efforts have resulted from the need to bypass the subjectivity and unrepeatability of traditional image-based classification methods, based on manual and expert interpretation (namely the manual digitisation of patterns in the acoustic images), improving the timing of such operations (keeping up with policy-making) and allow for repeatability; a crucial factor in further monitoring applications (e.g. Snellen et al. 2018). Current image-based classification approaches can be subdivided into two categories: unsupervised and supervised classification.

In unsupervised-type classification, regularities/homogeneities are searched in the unclassified data via clustering techniques. Under this approach, there is no a-priori information about the class nature (label) and statistically homogenous groupings (of pixels and/or image-objects) are the target of the clustering. The most prominent issue in unsupervised classification is the determination of the optimal number of clusters (e.g. Snellen et al., 2018). Several approaches have been proposed to tackle this issue. For example, Milligan and Cooper (1985) proposed and tested a set of thirty criteria to search the number of clusters objectively. In seafloor substrate mapping studies, this issue is exacerbated by the intrinsically noisy nature of the backscatter data, causing a natural overlap of the classes. Furthermore, unsupervised clustering often requires to initially map a large number of clusters and subsequently recur to aggregation via expert judgment of the groupings, limiting the automation and the repeatability and enhancing the subjectivity (e.g. Lathrop et al., 2006). It is however of interest for seafloor mappers to test the widest possible array of classifiers given that a consensus on the optimal implementable routine currently does not exist. Unlike an unsupervised method, where no a-priori information about the class labels is provided to the algorithm, supervised classification uses ground-truth information to train and test the classification results. The training process refers to the estimation of the set of parameters based on which the classifier can identify and label unseen data. The algorithm is therefore able to learn the patterns in the data based on a set of user predefined classes. Several supervised approaches (commonly referred to as Machine Learning approaches) have been proposed in the seafloor mapping literature over the past decade, including for example Maximum Likelihood Classification (e.g. Ierodiaconou et al., 2011), Support Vector Machines (e.g. Janowski et al., 2018), Artificial Neural Networks (e.g. Marsh and Brown, 2009) and Tree-based classifiers (such as Classification and Regression Trees [CART] and Random Forests) (e.g. Hasan et al., 2014).

The focus of this study was to merge and interpret acoustic surveys on the BPNS, acquired during eight oceanographic campaigns over the time span 2015-2018. Concurrent ground-truth samples were taken to appraise relationships between sediment grain-size variables and backscatter and to carry out a predictive surficial substrate mapping exercise. Samples were described in terms of a EUNIS level III (Evans et al. 2016) classes, and a finer-detail eight-class Folk classification scheme (Folk, 1954 - note that both schemes are based on Folk-type categories). Unsupervised k-means clustering and supervised Random Forest machine learning classifiers were used to predict the distribution of the substrate categories identified in the ground-truth data and tested for their predictive accuracy. Overall, this study aims to advance regional expertise in state-of-the-art seafloor mapping by multibeam echosounders. The novelty of this approach raises several scientific challenges and consequently, the aims of this investigation are manifold:

(1) Compile the multi-source multibeam survey data into a seamless backscatter and bathymetry dataset: this target promotes and advances the use of “natural MBES backscatter reference areas” as pragmatic at-sea solutions to control the backscatter measurements stability and repeatability, allowing merging of sparse yet overlapping surveys from different systems and platforms as well as subsequent change detection applications.

(2) Investigate the predictive accuracy of an unsupervised (k-means clustering) and a supervised (Random Forest machine learning) modelling/classification approach: this target advances the field of MBES based supervised seafloor classification by testing former- and current-generation classifiers. Due to the novelty of supervised learning approaches in acoustic seafloor classification, replication of investigations in a variety of environments and geographical settings is highly valuable to confirm methodologies and promote harmonisation across research teams and institutions.

(3) Investigate the predictive accuracy and backscatter discrimination potential for two habitat classification schemes (HCS), both based on Folk (1954) categories. A broad, regional-scale approach named EUNIS habitat level III (i.e. a simplified Folk scheme), and a finer-scale Folk scheme (i.e. conserving the original Folk classes and allowing for a higher diversity of sediment classes): this target investigates the “goodness of fit” of two important classification schemes used at the European level.

(4) Investigate relationships between backscatter and grain-size variables and discuss eventual mechanistic relationships (physical support) via the production of analytical plots. These have implications for assigning Folk classes to backscatter data: this aspect of the investigation attempts to

appraise how well can the backscatter measurements be related to readily accessible information of the benthic substrate from the ground-truth data. This is needed to advance the knowledge of the sediment acoustic relationships at sonar-operating-frequencies of interest, identifying limitations possibly having repercussions on classification schemes and the way we thematically map the seafloor.

(5) Apply modelling of angular response backscatter using the Generic Seafloor Acoustic Model to investigate appraise the physical differences in the thematic classes identified and predicted by the most accurate image-analysis method employed. The kind of modelling approach is highly flexible and allow fitting of experimental (field) backscatter data.

3.3 Materials and Methods

First, a general description of the study area is provided in terms of its geological, morpho-sedimentological, hydrodynamic and biological characteristics. Following, the planning, acquisition and processing of multibeam and ground-truth data measurements are discussed. Finally, analyses procedures are detailed, giving insights into the harmonization, exploration and integration of these datasets.

3.4 Study Area: Belgian Part of the North Sea

3.4.1 General seafloor setting

Part of the Greater North Sea and nested within the North-Western European Continental Shelf, the Belgian Part of the North Sea (hereafter BPNS) covers approximately 3600 km^2 ranging in depth (sensu m Lowest Astronomical Tide - LAT) from ~ 0 to -50 m. Three nautical miles (NM) delimitations denote a first depth zonation of this environment with 3, 12 and 24 NM marking coastal (nearshore ~ 0 to -15 m), central (mid shelf ~ -15 to -35 m respectively) and offshore (outer shelf ~ -35 to -50 m) continental shelf areas. Evidencing a significant sediment budget and hydrodynamic regime is the widespread presence of sandbanks and swales (relative to the coastline) which are the salient topographic features. These banks can be subdivided into four sub-groups with Coastal and Zeeland Banks distributed parallel to the coastline and the Hinder and Flemish Banks predominantly offsetting in respectively SW and NE direction (Figure 3.1a).

3.4.2 General geological background

The geological basement of the BPNS is characterized by solid Paleozoic layers covered by Cretaceous, Paleogene and Quaternary (Pleistocene and Holocene) sediments (Le Bot et al., 2003). The latter range in thickness from a few meters up to 50 m, with swales having a relatively thin cover (mostly less than 2.5 m) and sandbanks and scour hollows having a thicker cover and infill, respectively (10 to 50 m; see Le Bot et al., 2003; Hademenos et al., 2019 for a detailed account). The Holocene-originating sediments are mostly non-cemented and, except for relict gravel present in the central and offshore swales and planes (Van Lancker et al., 2007; De Clercq, 2018), their distribution and mobility is driven by hydrodynamic and climate regimes.

3.4.3 Morpho-sedimentological characterization

The present-day surficial sediments of the BPNS are primarily composed of siliciclastic sand. Sediments are coarser in the offshore aligning with the presence of the former Rhine-Meuse river valley. On a more local scale a progressive sorting can be observed from the gully to the flank and top of a sandbank, as well as along dune morphologies (Van Lancker, 1999). The sand fraction (0.0063-2 mm) is the predominant sediment type and mainly takes part in the present-day maintenance of sandbanks which are in turn populated by large and very-large dunes (sensu Ashley, 1990). As mentioned, coarser sands and gravel (> 2 mm and ranging from granule to boulders) are merely distributed in the offshore plains and within the swales where the Holocene cover is known to be less than 2.5 m (Deleu and Van Lancker, 2007), hence where the Paleogene clays potentially outcrop. Apart from these clayey outcrops, and the patchy occurrence of gravel, those deeper swales also allow deposition of fine-grained material. However, this silt enrichment is largely underestimated in modelled seabed maps (e.g., 0-1 % silt-clay content in the Hinder Bank swales, Verfaillie et al., 2006), partially due to sampling gear and observational constraints (Van Lancker et al., 2017).

3.4.4 Hydrodynamical setting

The tidal regime in the BPNS is characterised by semi-diurnal tides with mean ranges of ~4.3 m and ~2.8 during spring and neap tide respectively (referenced to Zeebrugge) (Van Lancker, 1999). The tidal ellipse is mainly oriented in SW-NE direction and the residual transport of the current is mainly oriented towards the NE (Van den Eynde, 2004). Current velocities during spring tide reach over 1 ms^{-1}. Higher current velocities have been recorded in the Westerschelde estuary and in the northern offshore areas, approaching the main channel of the Southern Bight of the North Sea (Lanckneus et al., 2001). High currents are generally associated to the incoming flood tide. However, exceptions exist in the swales of the regions of the Hinder and Flemish Banks where the peak in current velocity has been recorded

during the ebb flow, oriented in SW direction (Van Lancker and Verfaillie, 2005). Besides the tidal forcing, the BPNS experiences considerable wind and wave action. Winds originate predominantly from the SW remaining below 5 Beaufort for 90 % of the times (Verfaillie, 2008). Highest significant wave heights of more than 2 m are measured in the West Hinder area. Throughout the BPNS, the water-mass is year-round well-mixed, and no stratification is expected (van Leeuwen et al., 2015). Three main areas can be identified based on the magnitude of sediment transport during spring tide. In the offshore areas, beyond the 12 nautical miles delimitation, sediment transport is estimated < 0.5 tonnes $m^{-1}d^{-1}$; in the SW nearshore area of the BPNS between 0.5 and 1 tonnes $m^{-1}d^{-1,}$ whereas in the nearshore SE, where the turbidity maximum area is present, sediment transport is expected to exceed 1 tonne $m^{-1}d^{-1}$ (Lanckneus et al., 2001).

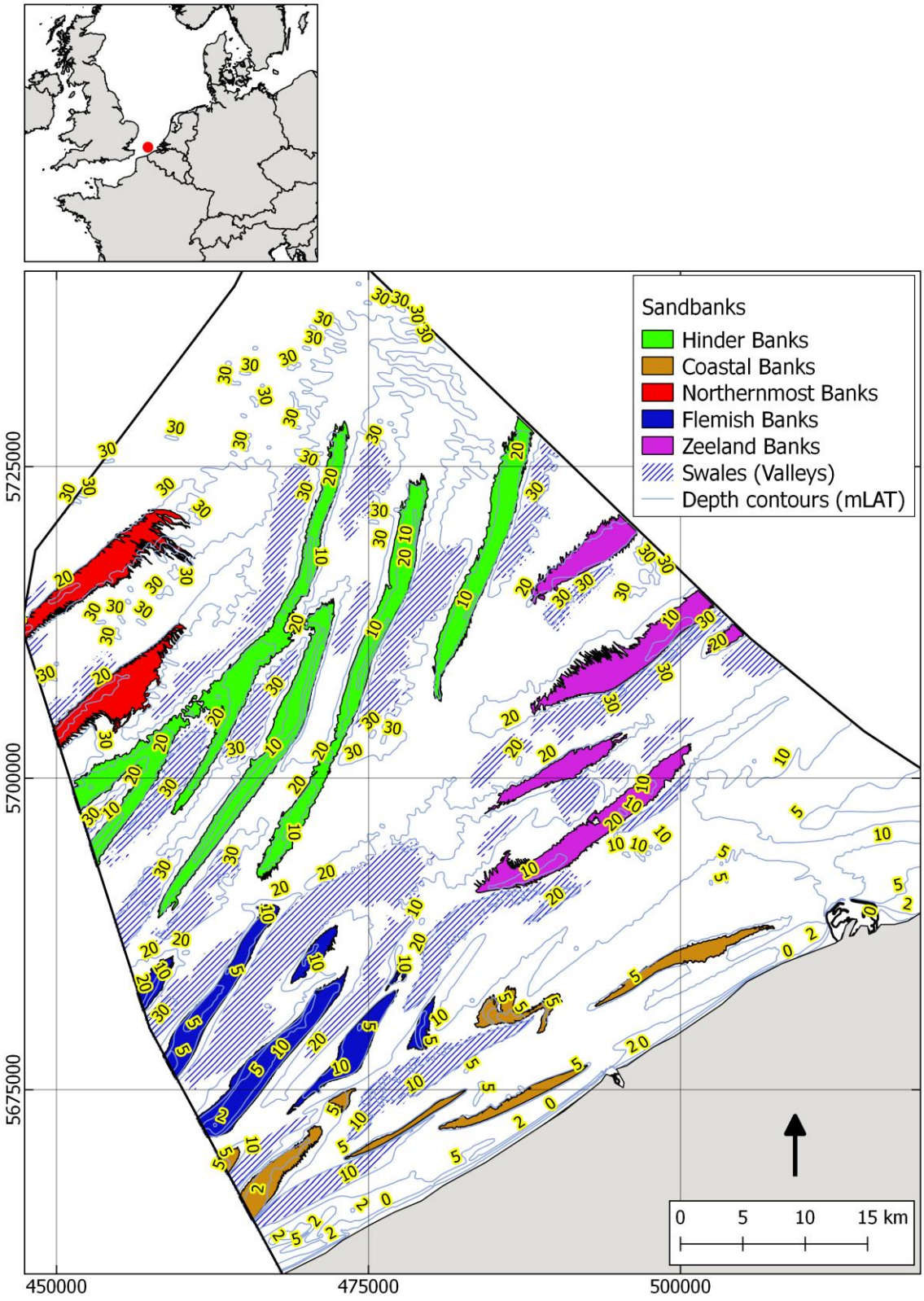


Figure 3.1a – Disposition and names of sandbanks and the geomorphologically most expressed swales in the Belgian Part of the North Sea. Sandbank and swales shapefiles used in the image: courtesy of MSc Lars Kint (RBINS) – geomorphometric analysis based on the 20 x 20 m MBES and SBES based legacy bathymetric dataset from https://www.afdelingkust.be/en/bathymetric-database. Depth contour lines from: Agency for Maritime Services and Coast – Coastal Division – Flemish hydrography.

3.4.5 Biological setting

The biological setting of the BPNS relates to five main soft-bottom subtidal macrobenthic communities which have been extensively explored (i.e. Degraer, 1999; Van Hoey et al., 2004; Degraer et al., 2008) whereas the hard substrate epibenthic communities have remained largely unexplored with the exception of early 1900's pioneering studies and some recent reiterations pertaining their potential distribution (Gilson, 1907; Kerckhof and Houziaux, 2003; Houziaux et al., 2008, 2011 and references therein). Presence of reef building polychaetes, namely *Lanice conchilega* and *Owenia fusiformis*, have been achieved through the use of very-high resolution side scan sonar surveys (Degraer et al., 2008) and very-high resolution multibeam bathymetry (Van Lancker et al., 2017). Regarding the soft-bottom (i.e. unconsolidated sediments) communities, the bivalve *Macoma baltica* community is characteristic of nearshore areas where fine sand and silt/muddy sediments are predominant, for example in the proximity of estuarine waters (i.e. Scheldt estuary and NE of the BPNS). Similarly, the *Abra alba* community is associated with nearshore sandy and muddy sediments (favouring a higher mud content and fine sands). A third community is that of the *Nephtys cirrosa* which, compared to the previously mentioned, has a low species diversity and abundance (unlike this polychaete species, bivalves come in dense aggregations). This community is typical of well-sorted sandy areas. The fourth community is that of the *Ophelia limacina* which generally favours coarser sands (but can also be found in areas of fine and medium sands) and is commonly associated with the presence of gravel and/or bioclastic detritus (i.e. shell debris). The *Barnea candida* community is the least represented given its rare habitat, i.e., compact Paleogene outcropping clay. Breine et al. (2018) recently acknowledged the establishment of a novel community, that of the *Magelona-Ensis leei* which is predominantly characterised by the non-indigenous bivalve species *Ensis leei*. This community is found across the entire coastal belt though it is distributed predominantly in very shallow waters and toward the NW portion of the coastal area where the mud/clay content is higher.

Hard substrates are generally recognised as promoting a rich and diverse occupancy by macrobenthic and benthic biota (McArthur et al., 2010; Taylor and Wilson, 2002). Due to the BPNS being a particularly challenging operational environment (from an historical marine-science research perspective), hard substrates have been poorly studied and remain thus far largely unexplored both in terms of their spatial distribution and consequently, their biology and status (Kerckhof and Houziaux, 2003). The majority of biological studies on gravel have been conducted in relation to artificial structures (e.g. Van Der Ben et al., 1977 in the intertidal belt, and Degraer et al., 2006, 2018 in the offshore windmills), and in the coastal subtidal area (in the vicinity of outcropping compact Paleogene clay (Degraer, 1999 and references therein). Historically, research on offshore gravel (pebble and boulder) fields associated with rich epifaunal communities, have been carried out by Gilson (1907) along the Hinder Banks, and were recently reiterated in Houziaux et al. (2008 and

references therein, 2011). Houziaux et al. (2011) reports on the historical (mid 1800's – early 1900) and present-day pervasive anthropic impact of fisheries in this area. Early efforts in mapping gravel in the Hinder bank region were presented by Veenstra (1969). More recently they were adjourned and extended by Van Lancker et al. (2007) and Van Lancker et al. (2016). The associated geological setting and origin of the gravel has been further studied by Deleu and Van Lancker (2007), Mathys (2009) and De Clercq (2018). Areas where gravel is most likely to be found are the swales situated between sandbanks, in the northernmost planes, as well as sporadically in the SW coastal area, where, in accordance to Kerckhof and Houziaux (2003), rich and diverse macrobenthic communities as those encounterable in French coastal waters (Davoult et al., 1988) may be found. The far offshore northernmost planes remain the least studied areas.

3.5 Multibeam survey strategy

Due to this area of the North Sea being a challenging environment from an oceanographic research perspective (i.e. owing to intense vessel traffic, routed navigation, widespread infrastructure, obstacles to navigation and frequently prohibitive weather conditions), surveys were strategically designed to accommodate the following:

(1) Over the BPNS (i.e. covering near- to offshore regions): encompassing the largest variety possible of morphosedimentary and hydrodynamic zones, covering the distribution of the main sediment classes as identified in the currently available sediment distribution maps for this region.

(2) Within the Habitat Directive area (Flemish Banks area) and specifically to the "Fisheries zones" (Z3 and Z4 of the former Marine Spatial Plan (Belgian Royal Decree of March 20, 2014 – see e.g. Olsen et al., 2014) of the Hinder bank region: targeting ecologically noteworthy areas historically known to be characterized by gravel and diverse assemblages of benthic biota (Gilson, 1907; Deleu and Van Lancker, 2007; Houziaux et al., 2011; Van Lancker et al., 2017).

(3) Survey areas in the far and near-field of anthropogenic activities (e.g. disposal grounds of dredged material, aggregate extraction and fisheries areas) where seafloor changes are expected to occur and future monitoring targeted.

Figure 3.1b provides a cartographic summary of the surveys conducted in respect to some of the major anthropogenic activities and superimposed on the current distribution of the European Nature Information System (EUNIS) level III sediment classes (i.e. sediment type described by a simplified Folk scheme accounting for three main categories; sandy mud to mud, muddy sand to sand, and coarse sediments). Such maps are currently available at scales of 1:250.000 and 1:100.000. They can be downloaded from the European Marine Observation Data Network

[EMODNET] https://www.emodnet-geology.eu/map-viewer/?p=seabed_substrate (see Kaskela et al., 2019).

The newly-acquired acoustic surveys sum up to approximately 150 km^2 (about 4 % of the BPNS) and are composed of a combination of “full-coverage” and “transect” surveys. Full-coverage surveys consist in a set of parallel navigation lines oriented in accordance to the local main tidal axis; these result in a full-coverage map of an area with a 20 % overlap between adjacent lines. Alternatively, the transect approach consists of “reconnaissance” surveys acquired along trajectories (e.g. the EM2040-S survey is a ~200 km long trajectory parallel to the coastline; Fig. 1b). Parallel lines were equally navigated along the trajectories to allow enough overlap enabling comparison with future surveys. Data were also logged while transiting from Oostende harbour towards the Hinder Bank region (Fig. 3.1b).

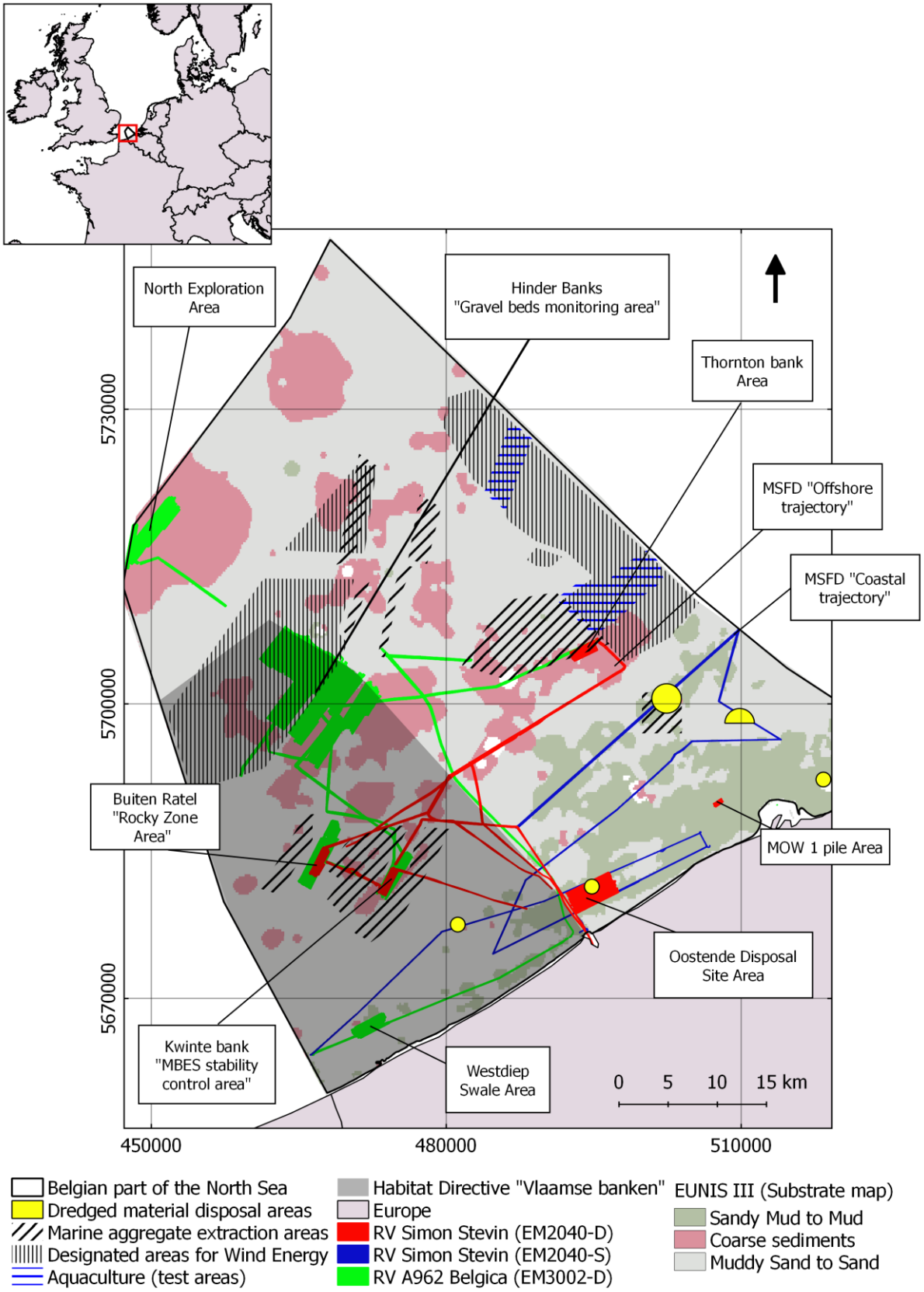


Figure 3.1b – Map of the MBES surveys conducted between 2015 and 2018 in the Belgian Part of the North Sea. Caption continues on next page.

The upper quadrant shows the location of this study area within the Southern North Sea. Green, red and blue filled polygons show the extent of the EM3002D, EM2040D and EM2040S MBES data respectively. Some of the anthropogenic activities are denoted by patterned polygons and include: Disposal grounds of dredged material (yellow filled), areas of marine aggregate extraction (blue thin line pattern), envisaged aquaculture management zones (Blue thin line pattern) and designated areas for windfarms (existing and planned - black thin line pattern). The Habitat Directive area is denoted by a grey shaded polygon. Data are superimposed on the Initial Assessment map of the EUNIS III substrate classes of the Belgian Part of the North Sea (see Van Lancker et al. (2018) for a history on the mapping procedure). All data projected in WGS 84, UTM 31 N.

3.5.1 MBES data acquisition and processing

High-frequency (300 kHz) multibeam surveys were conducted over the course of three years (2015-2018) covering nearshore to offshore areas of the BPNS. Kongsberg Maritime systems EM3002 dual and EM2040 installed on RV A962 *Belgica* and *Simon Stevin*, respectively, were operated during eight oceanographic campaigns. Data were logged using the Kongsberg Maritime's acquisition software SIS. Both echosounders were operated in high-density equidistant mode, forming 508 (1.5° x 1.5°) and 800 (1° x 1°) soundings per ping, respectively for the EM3002 and EM2040 dual systems. Real-time corrections for sound velocity in the water column were obtained by a Valeport mini-SVS sensor installed in proximity of the transducers. Precise positioning and vessel motion data (roll, yaw, pitch and heave) were recorded by an MGB Tech with Septentrio AsteRx2eH RTK heading receiver and a Seatex MRU 5 unit for the EM3002D, and by an MGB Tech with Septentrio AsteRx2eL RTK heading receiver and a XBlue Octans motion sensor for the EM2040D. The EM2040 on RV *Simon Stevin* was upgraded to a dual system during 2017, whereas the EM3002D remained unchanged throughout the time span of acquisition. Figure 3.2 shows the surveyed areas extents, summarised by campaign, and Table 3.1 provides details of the acquisitions relevant for the subsequent inter-calibration of these datasets to produce maps of "seamless" backscatter and bathymetry.

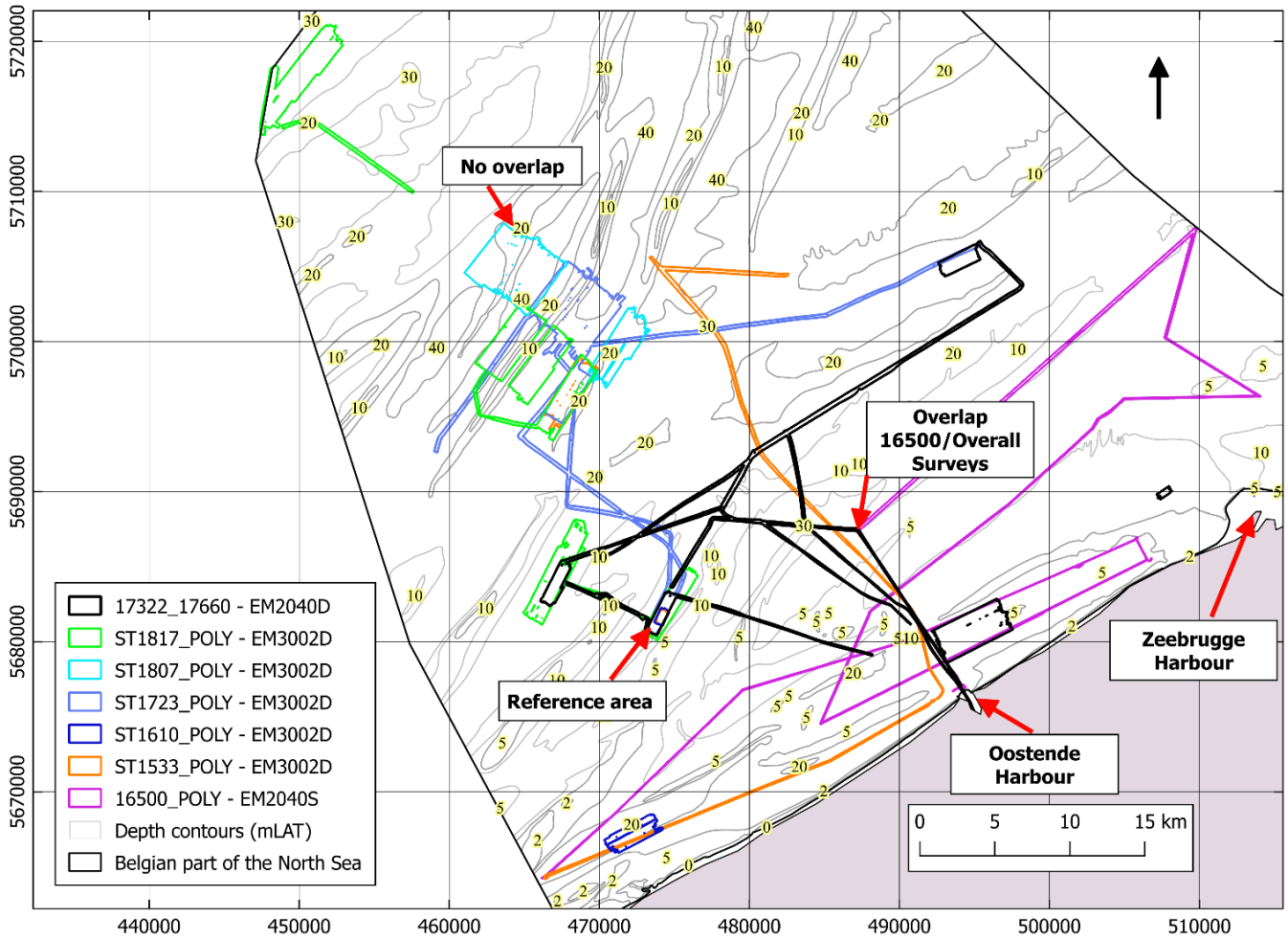


Figure 3.2 – Overview of the MBES surveys summarised by campaign. Areas of overlap between surveys are reported. The campaign codes reported in the legend relate to the campaign codes and corresponding campaign reports that can be queried from http://www.vliz.be/vmdcdata/midas/cruise.php?showcruise=1 (for RV *Simon Stevin*) and from https://odnature.naturalsciences.be/*Belgica*/en/ (for RV *Belgica*). All data projected in WGS 84, UTM 31 N.

Table 3.1 – Overview of the oceanographic campaigns. The dB offsets from the nominal truth (i.e. the ST1533 campaign) used to produce the seamless backscatter map are reported in the last column. α_w refers to the absorption coefficient correction (sensu Francois and Garrison, 1982a, b).

Sonar	RV	Campaign code (Fig. 3.2)	Pulse length (µs)	Freq. (kHz)	Year	Month	Overlap	α_w (corr.)	Offset (dB)
EM3002-D	*Belgica*	ST1533	150	300	2015	Dec.	y	y	Ref.
EM3002-D	*Belgica*	ST1610	150	300	2016	Mar.	y	y	< 1 dB
EM3002-D	*Belgica*	ST1723	200	300	2017	Jul.	y	y	< 2 dB
EM3002-D	*Belgica*	ST1807	200	300	2018	Mar.	na	y	na
EM3002-D	*Belgica*	ST1817	200	300	2018	Jul.	y	y	< 1 dB
EM2040-S	*Simon Stevin*	16-500	216	300	2016	Jun.	y	y	< 1.5 dB
EM2040-D	*Simon Stevin*	17-322	288	300	2017	May	y	y	< 2 dB
EM2040-D	*Simon Stevin*	17-660	101	300	2017	Nov.	y	y	< 2 dB

Bathymetry data processing was carried out using QPS Qimera© (v1.2.4.429a). Real-time kinematic (RTK) and GPS modelled tide solutions were used to correct the real time navigation data. In turn, manual edits were applied to the soundings, referenced to the Lowest Astronomical Tide datum (WGS 84, UTM 31N). Data were gridded to a 5 m horizontal resolution (Fig. 3.3).

Backscatter data processing was carried out in QPS Fledermaus Geocoder© (FMGT) software (v7.4.5. b). To allow data inter-comparison, a strictly standardised procedure was maintained during the processing phase. FMGT mosaic processing parameters (“pipeline settings”) were maintained as close as possible to the default settings of both echosounder models. All beams from the “beam time series” (from the Kongsberg datagram) were kept. Absorption in the water column was compensated by the absorption coefficients (sensu Francois and Garrison, 1982b, 1982a) in the raw datagram. This coefficient was updated every 30 minutes while logging the data and computed according to the local surface seawater properties. The necessary water-medium parameters were obtained by the On-Board Data Acquisition System (ODAS), logging these data at 10-s intervals. In this region of the North Sea, surface water values may be considered sufficiently representative of the whole water column (van Leeuwen et al., 2015). using FMGT, the angular dependence was compensated leaving parameters as close as possible to the default settings i.e. an Angular Varied Gain window size of 300 pings and the default “mosaic processing” settings. the sole modification was the average reference angle used to normalize the data, set in the range 43°- 47°. The true ensonified area was accounted for by inclusion of a Digital Terrain Model (DTM) in the processing.

To allow merging of the disparate backscatter datasets and produce a seamless map of reflectivity, a methodology similarly to Hughes-Clarke et al., (2008) and Misiuk et al., (2018) was applied. The approach may be referred to as a “cross-

calibration propagation" and consists of selecting a reference survey and adjusting all other surveys by overlap to the nominal truth (dB offsets are reported in Table 3.1). Here the overlap refers to a highly stable relative calibration reference area (Kwinte bank swale) for which a detailed description can be found in (Roche et al., 2018). The advantage of surveying a stable reference area in combination to a rigorous standard in acquisition and processing of the measurements, is that the data stability can be controlled and subsequently the repeatability of the measurements guaranteed. Except for the campaign 16-500 using EM2040 single system and campaign ST1807 using the EM3002D, each survey launch saw the acquisition of data over the reference area based on which empirical dB offsets could be derived to correct the data to a nominal truth survey (Table 3.1). For the 16-500 launch, an area of overlap on a flat and sandy substrate allowed deriving the shift value whereas for the ST1807 campaign no overlap was achieved. The resulting backscatter map is displayed in Figure 3.4 showing the seamless character achieved after applying the offsets and merging the surveys.

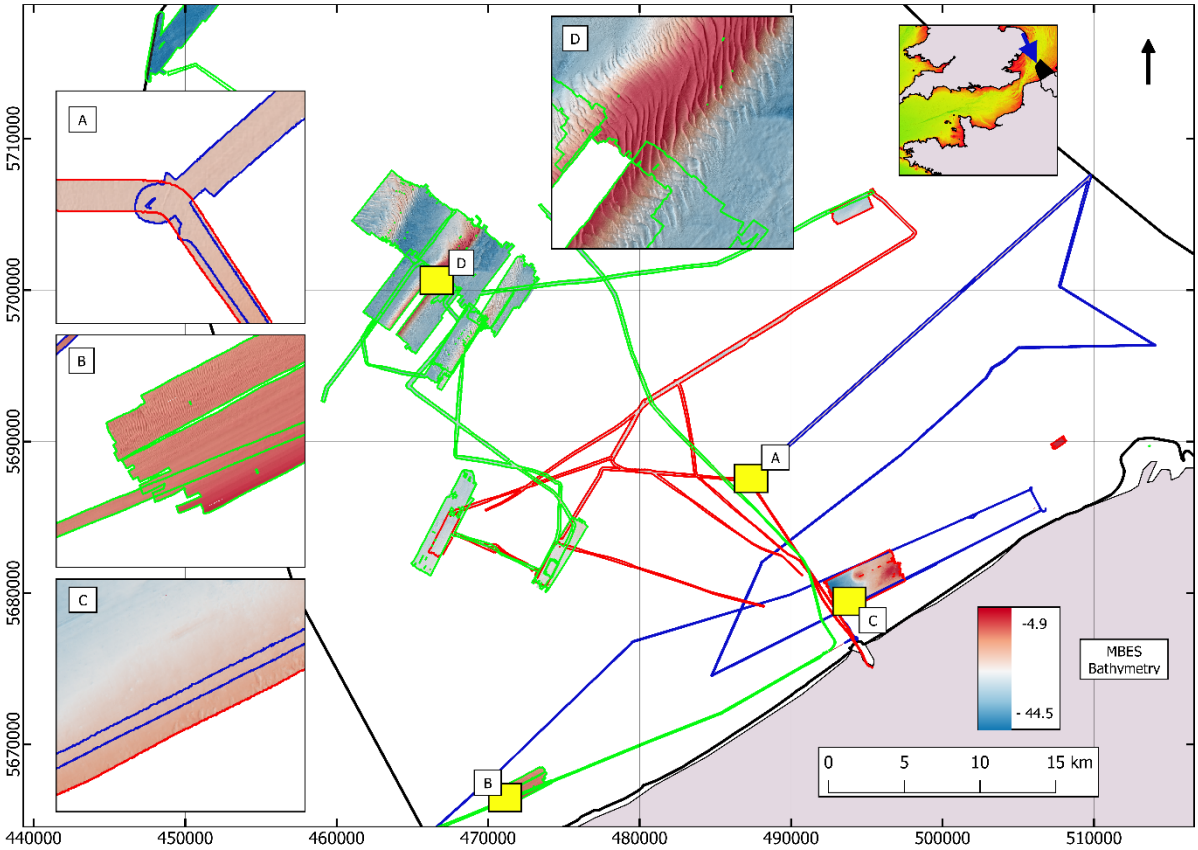


Figure 3.3 – Processed, cleaned and merged bathymetric dataset. Quadrants A, B, C and D show details of some areas of intersection between surveys showing the good seamless character (no bathymetric differences producing "step-like" artefacts). All data projected in WGS 84, UTM 31 N. Polygons are colour-coded as in Fig. 3.1b.

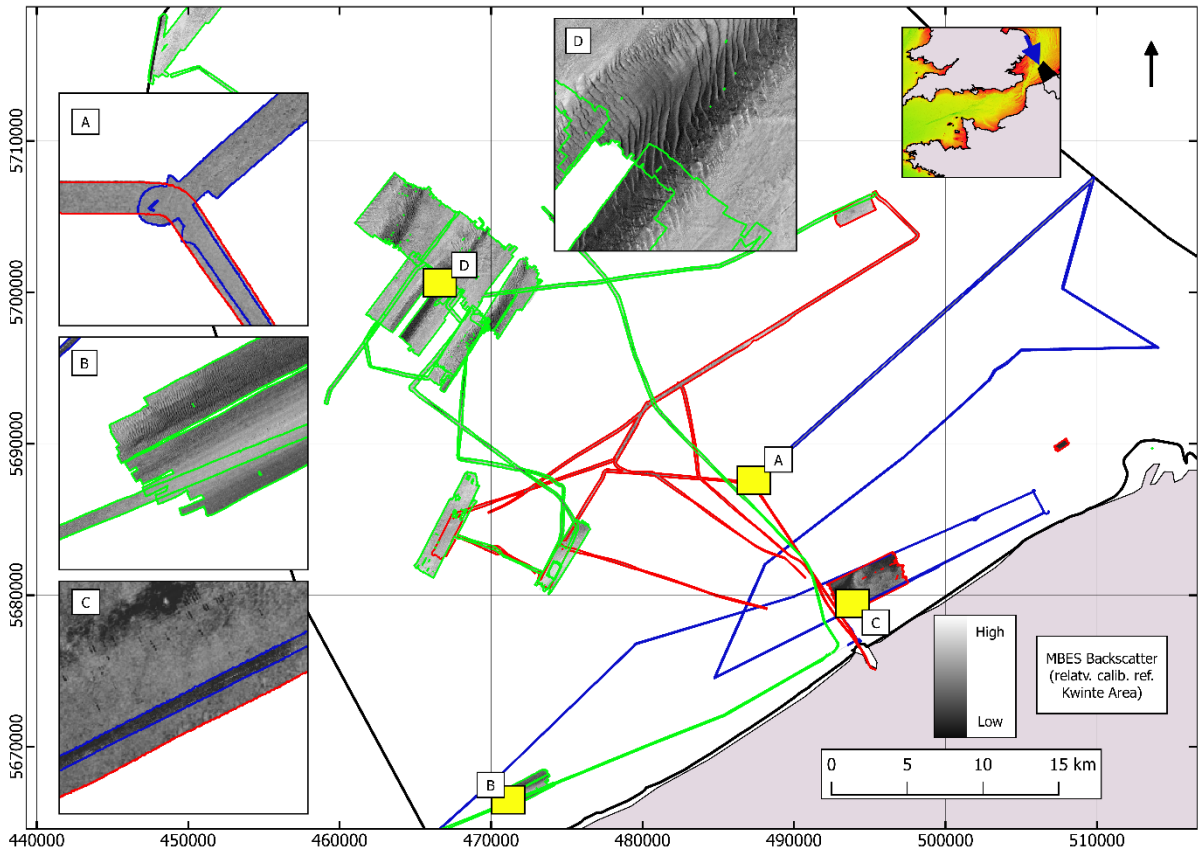


Figure 3.4 – Merged multi-source multibeam backscatter mosaics. Quadrants A, B and C detail the seamless character of backscatter obtained by different echosounders and platforms. A) Detail of the seamless character achieved between EM3002-D and EM2040-S echosounders. B) Details of the seamless character achieved between different surveys of the EM3002-D on RV *Belgica*. C) Sedimentary changes between EM2040-S (blue outlined polygon) and EM2040-D (red outlined polygon) related to the harbour and channel maintenance anthropic activities of sediment dredging and disposal in the designated dumping site in the vicinity of Oostende harbour (Lauwaert et al. 2016). D) Seamless character of EM3002-D surveys within the Hinder bank survey area. All data projected in WGS 84, UTM 31 N. Polygons are colour-coded as in Fig. 3.1b.

The cross-calibration propagation was carried out by applying the dB empirical offsets directly to the mosaicked compensated backscatter grids (an approach aligning to that of Urgeles et al. (2002) and Hughes-Clarke et al. (2008)). Considering the processing chain of FMGT (Lamarche and Lurton, 2018; Schimel et al., 2018), it is assumed that all angle dependencies have been compensated for, i.e. those caused by the MBES directivity pattern and those from the backscatter angular dependence. Therefore, the mosaic is representative of the average backscatter strength of the seafloor normalised to a conventional average reference angle in the range 43°- 47° (namely BS_{45}), including the systems sensitivity (i.e. ± 1 dB for the Kongsberg systems herein used). As such, the dB offsets represent average shifts at 45° and by referring all surveys and sounders to the same nominal truth, the sounder sensitivity is corrected for, leaving only the seamless character (continuity of acoustic facies) of the average response; lawful for a regional compilation of backscatter maps.

3.5.2 Modelling of angular response backscatter

Besides processing of the compensated backscatter imagery for further classification, SonarScope® software (IFREMER; Augustin and Lurton, 2005; Lamarche et al. 2011) was used to compute angular response curves from the beam intensity datagrams within an insonified area across the swath covering a portion of the seabed around each ground-truth station. The insonified area considered in each swath is proportional to the water depth at the sample position. This implied considering about fifty pings for 10 m depths (coastal surveys) and about twenty pings at 40 m depths (mid-shelf and offshore surveys). The assumption of seafloor homogeneity inside the insonified area was put forward although this might have been violated in the coastal area where considerable patchiness is expected. In order to retrieve the best estimate of the raw backscatter angular response from the Kongsberg Beam Intensity datagrams, the processing chain as in Roche et al. (2018), used for example in Lamarche et al (2011), Fezzani and Berger (2018) and Montereale-Gavazzi et al. (2019), was followed. It entails four main steps: 1) correction for sound absorption computed from the Levitus salinity database and the sound velocity profile in the SonarScope built-in facilities (Levitus et al., 1994; Levitus and Boyer, 1994); 2) correction of the ensonified area using the real incidence angle as from the tide-corrected bathymetric models; 3) removal of all angle-dependent corrections introduced by the sonar manufacturer, and 4) computation of the AR curves within the insonified area across the swath.

In a following step, the raw curves were fitted with the Generic Seafloor Acoustic Backscatter (GSAB) model. This model describes the angular response using a combination of three statistical angle dependency laws: a Gaussian law fits the specular region of the angular range, whereas a Lambert-like law is used to fit the oblique and fall-off regions; a second Gaussian law fills the intersection between the former components. Overall, four to six parameters (A-F - Equation below) are used to describe the angular response in terms of dB intensities and angular extents, thus the behaviour of backscatter as a function of incidence angle for a variety of sediment types. Figure 3.5 provides an example of fitting GSAB parameters to measured data. The parameters used to fit the measured angular response do not directly relate to geological and geotechnical sediment properties as in the geoacoustical backscatter models (i.e. Jackson et al., 1986), though they provide a physical description of the backscattering from seafloor sediments.

$$BS(\theta) = 10\log\left[A\exp(-\theta^2/2B^2) + C\cos^D\theta + E\exp(-\theta^2/2F^2)\right]$$

For example, the parameter A is related to the specular coherent reflection (the maximal amplitude intensity in dB) and it will be highest for smooth and fluid-like sediments and for strong contrast in acoustic impedances between the water medium and the seafloor interface. B refers to the angular extent of the specular domain and relates to the interface roughness. The parameter C relates to the mean backscatter level (in dB) in the oblique domain of the angular range. This parameter

is associated to the Lambert's law which describes the backscattering phenomenon at oblique angles for rough and coarse water-sediment interfaces (at a roughness scale comparable to the acoustical wavelength ~ 5 mm at 300 kHz). C is also dependent on the volume, thus on the in-sediment heterogeneity. It is found to increase with increasing roughness and impedance as well as in the presence and characteristics of buried scatterers, possibly being the dominant scattering mechanisms in soft sediments. The parameter D is the decay rate of the backscatter strength (in dB) with grazing angle and is found to increase for soft and flat seafloor interfaces. Without clear physical meaning, the parameter E describes the maximum level (in dB) of the transitory region between the specular and grazing angular domains whereas F refers to its angular extent (Lamarche et al., 2011).

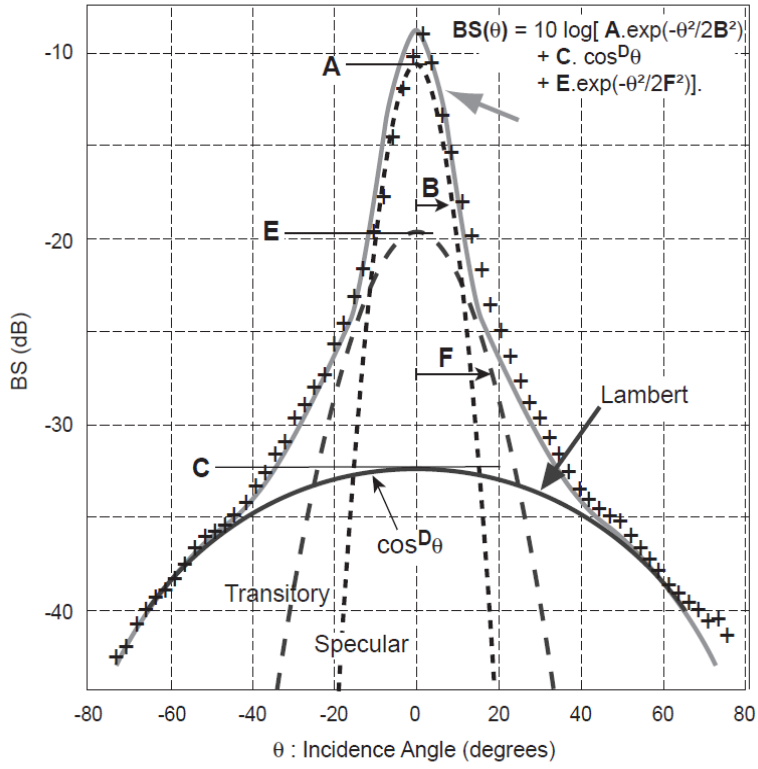


Figure 3.5 – Fitting of GSAB model (grey solid line) to measured and corrected raw backscatter data (crosses) via a combination of Gaussian and Lambert-like functions (dashed and solid black lines) adjusting parameters A to F. The adjustment of the parameters is carried out iteratively until an optimal fit to the measured data is found (from Lamarche et al., 2011). The model will find a symmetric curve, disregarding the potential differences in sediment type detected at port and starboard track sides.

Modelled average angular response curves for each of the 163 sample locations were in turn related to the most discriminative Folk classification scheme identified (see next section and Results – Data exploration). Due to the echosounder systems not being effectively intercalibrated (a simple cross-calibration propagation was applied only to the compensated backscatter imagery - see previous section), averaged GSAB curves for the classes identified were plotted by echosounder type, hence for the EM2040-S and Dual and for the EM3002D. Furthermore, a set of

samples were selected ad-hoc based on the homogeneity of the backscatter mosaic at their location to exemplify the potential of AR in capturing physical differences in substrate composition for the classes identified.

Unfortunately, the RV *Belgica* EM3002D and RV *Simon Stevin* EM2040S and EM2040D MBES are currently not intercalibrated, so the BS angular response curves resulting from these different MBES are not directly comparable in terms of dB values. While this remains a serious limitation, an ongoing project in collaboration with IFREMER laboratory underwater acoustics and the Continental Shelf Service of Belgium is currently being dedicated to this issue with an approach combining MBES measurement and fully-calibrated SBES measurements on a set of reference areas (details of these areas and approach can be found in Roche et al. (2018) and Eleftherakis et al. (2018) respectively).

3.5.3 Ground-truth data acquisition and processing

Collection of concomitant ground-truth data is necessary to validate the nature of the acoustic data and ultimately to derive confidence metrics expressing the validity (i.e. visual and statistical accuracy assessments) of the thematic models produced. The ground-truth data herein used were acquired in complement to each MBES survey (i.e. within ~48 hours from the acoustic survey completion) and are therefore closely representative of the seafloor spatio-temporal status at the time of each survey. The sampling effort was planned in such way that it was representative of the area (i.e. backscatter map) being sampled. To achieve this, the backscatter cumulative distribution of the study area was visually compared to that of the hypothesised sample locations (as in Montereale-Gavazzi et al., 2017) prior to the sampling action. Several gears were tested and deployed including physical (i.e. grab and core sampling) and optical (videography (drop-videoframe) and sediment profile imaging (SPI) instrumentation) (Figure 3.6). The choice of gear was largely dictated by the expected type of substrate being sampled. For example, the Hamon grab sampler and video observations were the instruments of choice within the gravel areas (where Van Veen and box core systematically fail), whereas box cores, Van Veen grabs and SPI were favoured within the soft sediment areas. Grab samples are the focus of the present investigation, however videography and photography were very useful and assisted the qualitative interpretation of the maps produced. To understand backscatter variation in the patchy gravel areas, videography of a larger environment is critical and is work in progress.

Only samples overlapping with the acoustic surveys were kept for further analysis. These sum up to an overall n = 163 samples; subsequently described in terms of surficial substrate type. Sample coordinates were geo-referenced and automatically corrected during the acquisition for the DGPS antenna layback accounting for the main source of positional error. Samples were described combining visual and expert observations (recorded on board) with grain-size parameters derived by

sediment analysis using a Malvern Mastersizer 3000 instrument (following the standard sample preparation routine as in Montereale-Gavazzi et al., 2017). The results of the laser granulometry were processed in GRADISTAT (Blott and Pye, 2001) from which metric sample statistics were kept for further analysis. Since only the fraction ≤ 1 mm could be analysed by the Malvern, the percentage of the coarse fraction (> 1 mm, namely bioclastic detritus and gravel) was visually estimated by observation of photographs and notes taken on-board for every retrieved sample (thus scoring a qualitative gravel percentage). Samples were split according to a random stratified split rule (70 – 30 for training and validation subsets respectively).
Only features visible at the water-sediment interface were described (except for Hamon grab samples where the sampling does not preserve the vertical integrity of the seafloor) and summarised into thematic classes according to two classifications schemes commonly applied in the European underwater mapping context:

(1) A broad classification scheme: European Nature Information System (EUNIS) habitat level III classification (see for example: Galparsoro et al., 2012). For the BPNS, samples are then summarised into three predominant Folk-type substrate classes: sandy mud to mud (**sM** to **M**), muddy sand to sand (**mS** to **S**) and coarse sediments (**C**).

(2) A finer detail scheme, allowing for a more detailed distinction of sediment classes, based on the Folk (1954) classification where mapping is based on the relative proportions of the three-size fractions categorized into Mud (with a particle size diameter d < 63µm), Sand (d = > 63 µm < 2 mm) and Gravel (d = > 2 mm). Here, following an exploratory data analysis (see Results/ Data exploration) five substrate classes ranging from sandy mud (**sM**) to sandy gravel (**sG**) where selected for the mapping: sandy Mud (**sM**; ≤ 5% gravel with a sand:mud ratio between 1:9 and 1:1), muddy Sand (**mS**; ≤ 5% gravel with a sand:mud ratio between 1:1 and 9:1), Sand (**S**; ≤ 5% gravel and a sand:mud ratio of at least 9:1), gravelly Sand (**gS**; up to 30 % gravel and a sand:mud ratio of at least 9:1) and sandy Gravel (**sG**; up to 80% gravel and a sand:mud ratio of at least 9:1). In the sG class are also included samples with more than 80 % gravel due to their paucity. This generalises the description (thus subsequent prediction) of this class though it deals with the scarcity of samples for this class.

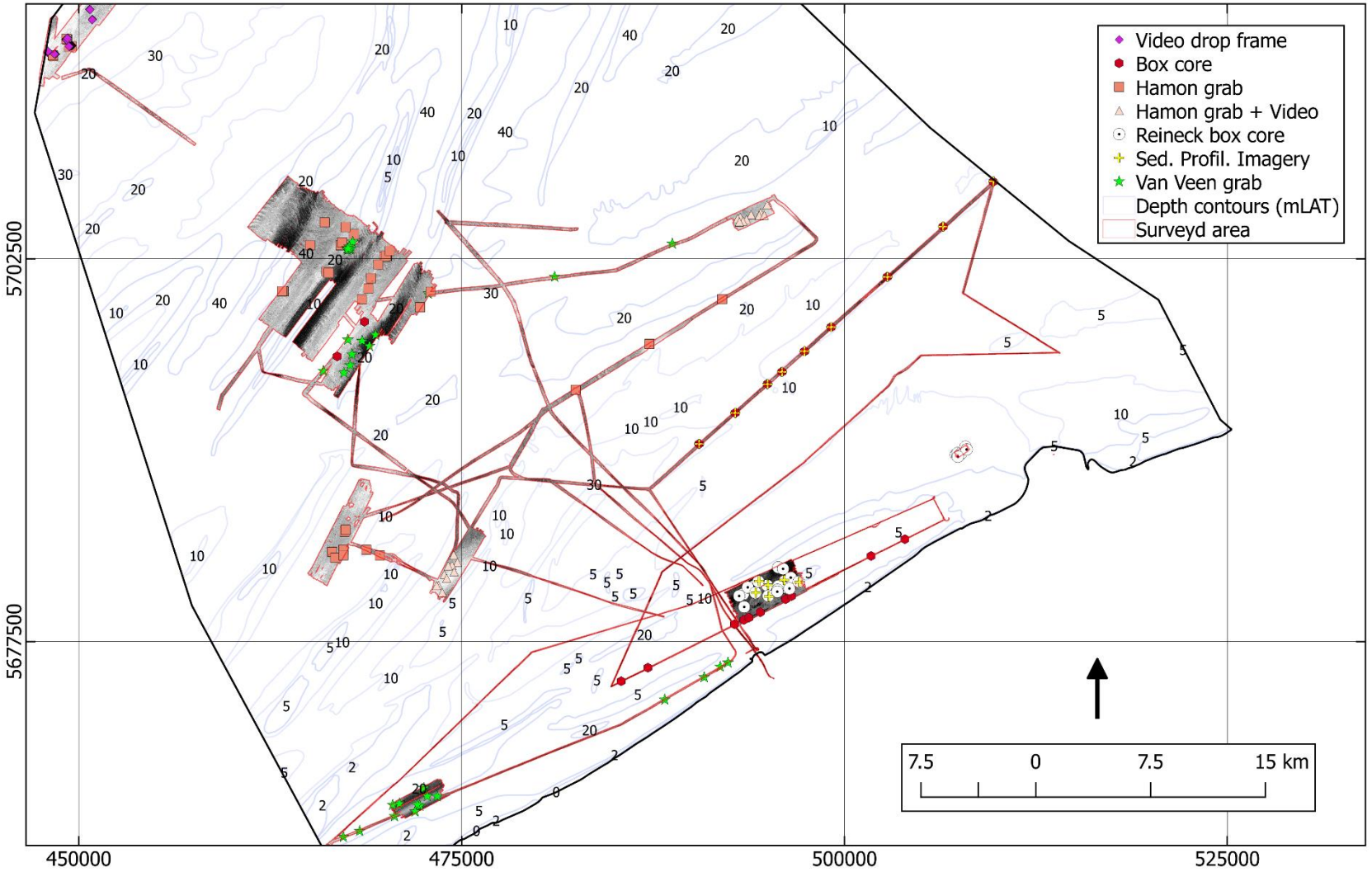


Figure 3.6 – Summary of the sampling effort summarised by gear and displayed over the entire MBES acquisition in the period 2015 – 2018. Contour lines are displayed in light grey and labelled with positive depth (mLAT) values. All data projected in WGS 84, UTM 31 N.

3.6 Exploratory data analysis: Relationships between backscatter and sediment type

Relationships between MBES backscatter and sediment type were initially investigated using boxplots summarising bathymetry and backscatter statistics grouped by EUNIS III and Folk sediment categories. This provides insights into the class separation potential (i.e. the discriminative power of the data in respect to the proposed classifications schemes). Cumulative distributions of backscatter and bathymetry were compared between the entire study area and training and validation sample sets to visually assess their representativeness, thus their viability for the ASC routines. Simple linear regression was used to assess relationships between the average backscatter extracted from a 25 m buffer around each sample location and the median grain-size diameter (D50). A more insightful analysis was based on relationships between percent weight of individual size fractions and mean backscatter. This was possible based on a set of 12 samples acquired within the Hinder bank and the Northern Exploration area (see Fig. 3.1B), where sandy and gravel areas predominate. They were dry sieved using mesh size intervals of 0.063, 0.125, 0.355, 0.5, 1, 2, 4, 8 and 10 mm. Here, the association of backscatter with the weight percentage of finer (i.e. range < 0.5 mm) and coarser fractions (i.e. range 1 – 10 mm), as well as the correlation between backscatter and the percentage weight of each sieve fraction were investigated.

3.7 Substrate modelling approach

Two modelling approaches were tested to predict class membership of both classification schemes over the full extent of the seamless 5 m datasets. Unsupervised clustering via k-means (sensu Hartigan and Wong, 1979), and supervised classification via Random Forest (RF; Breiman, 2001) were chosen. k-means clustering is amongst the most widely applied data clustering technique, including numerous examples in the marine literature (e.g. Hewitt et al., 2004; Fonseca and Calder, 2007; Alevizos et al., 2015, Snellen et al., 2018, Fezzani and Berger, 2018). Supervised Random Forest (RF) was selected to test the performance of backscatter in combination with bathymetry alone, as well as in combination with a set of derivatives of the primary MBES data (i.e. further explanatory/predictor variables). RF models have been reported to achieve high predictive accuracy in recent studies focusing on the comparison of supervised classifications of MBES data into substrate and habitat maps (e.g. Diesing et al., 2014; Diesing and Stephens, 2015; Ierodiaconou et al., 2018; Turner et al., 2018; Porskamp et al., 2018) and have generally proven highly successful in remote sensing applications (Belgiu and Drăguţ, 2016).

3.7.1 Unsupervised approach

k-means is an unsupervised clustering method that seeks to reduce the within cluster/group variance while maximising the between groups variance through an

iterative process of cluster centres assignment and re-allocation. When the parameter k is known (i.e. the number of classes), the algorithm can be executed in the following three main steps: 1) assignment of initial cluster centres (centroids of the proto clusters), 2) allocation of the data points to their closest cluster, and 3) iterative re-allocations of data points to the clusters for which the Euclidean distances (in the feature space) are smallest. A solution is found when all data points have been allocated. A random initialisation partition number was assigned prior to the clustering to allow for reproducibility of results. Besides the comparison with ground-truth data, the optimal number of clusters was searched via the implementation of two criteria as in Eleftherakis (2013): The Within Group Sum of Squared Distances (WGSSD – a measure of within cluster homogeneity) and the Silhouette coefficient of the k-means clusters (Rousseeuw, 1987). The first criterion is a metric that looks at the WGSSD as a function of the number of clusters. The optimum is chosen where adding clusters does not improve the WGSSD. The second metric quantifies the dissimilarity of single data points to the overall points of their cluster and returns measures in the range 0-1. A Silhouette coefficient > 0.5 is indicative of sufficient class separation whereas below this threshold classes are found to be significantly overlapping (Eleftherakis, 2013). Both criteria were tested running k-means for a 2 to 15 cluster solution. K-means classification approach was implemented in R (R Development Core Team, 2015) using the RStoolbox built-in functions. The nstart parameter (describing the number of attempts of initial centroid configurations), was set to 50 to allow the algorithm to identify an optimal initial allocation of centroids (i.e. centres of the proto-clusters) from a sufficient number of initial attempts. Labelling of identified clusters into sediment classes was achieved by sorting the cluster averages under the assumption of linearity between backscatter and sediment grain size (supported by the exploratory data analysis) and supported by expert interpretation. The optimal number of clusters was identified in the range of 3-5 by the WGSSD equally to the Silhouette method proposing the optimal solution at 3-5 clusters with 0.58 and 0.57 respectively.

3.7.2 Supervised approach

Random Forest (Breiman, 2001) is an ensemble algorithm based on the fundamental unit of machine learning: the decision tree. It can be used for regression (on numerical data) and for classification (on categorical data). It is referred to as an “ensemble” method as the ultimate classification prediction is based on the aggregation of majority votes obtained from several (forming a “forest”) randomly constructed decision trees (i.e. using a randomly sampled subset of the training samples with replacement - bootstrapped samples). Throughout the iterative process of “growing” trees, a part of the training samples is left out of the process (“out of bag” samples) and used to internally cross-validate the predictive performance of each tree. The underlying principle of this modelling approach is that the inherent tendency of single decision trees to overfit the predictions is overcome by bootstrapping the input training data and by aggregating (an approach referred to as

"bagging") the predictions of several randomly constructed trees, resulting in more stable decisions. This randomisation and aggregation approach at row and column levels has been found to effectively keep the bias of the training samples low while reducing the variance of the predictions (Breiman, 2001). Further randomness in the model is in the selection of variables tested at the splits of each tree; the contribution of each randomly selected descriptor to the overall classification error is estimated by this iterative approach (and by the a priori application of a variable selection routine). In turn, a majority voting is applied to decide the final class prediction from the various trees.

RF modelling was implemented in R (R Development Core Team, 2015) using the RandomForest (Liaw and Wiener, 2002) built-in functions. Using the same training and validation datasets, two RF models were trained:

1. Backscatter and bathymetry (hereafter referred to as RF_{Simple}) and
2. Backscatter, bathymetry and a set of selected backscatter and bathymetric derivatives (hereafter referred to as RF_{++}).

These choices were made in the view of testing whether the inclusion of bathymetry data alone would suffice to improve the predictive accuracy, or if this would require additional derivatives (i.e. explanatory variables).

3.7.3 MBES derivatives

To enhance the local characterisation of the primary MBES dataset and identify homogenous areas of substrate and morphology using the supervised approach, a set of secondary spatial derivatives (predictor variables/predictors) were produced from backscatter and bathymetric grids (Table 3.2; Figure 3.7). Selection of the spatial layers was based on their expected influence on the distribution of sediment type and due to their ability to enhance the predictive accuracy of seafloor substrate and benthic habitat thematic models in previous research (e.g. Lecours et al., 2015; Rattray et al., 2013; Ierodiaconou et al., 2018). All layers were computed within a 3 x 3 kernel.

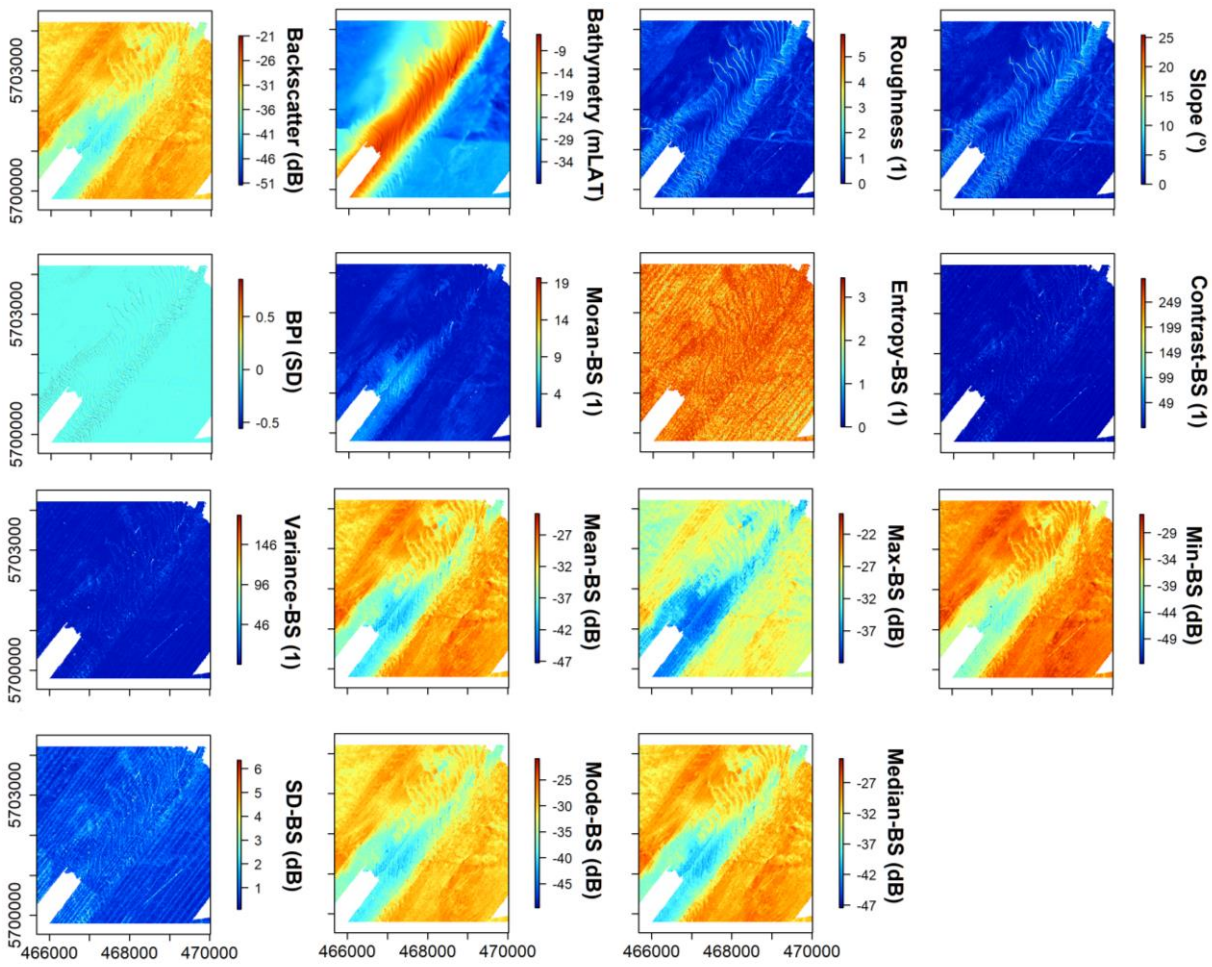


Figure 3.7 – Details of the MBES primary backscatter and bathymetry grids and the derivatives used in this investigation. All data projected in WGS 84, UTM 31 N. The displayed area is a detail of the Hinder bank area (West Hinder - see Fig. 1b). The unit (1) indicates scalar quantity.

Table 3.2 - Predictor variables derived from the primary MBES data.

Morphometric derivatives (from bathymetry)		Software
Slope	Maximum change in elevation between each cell and cells in its analysis neighbourhood (3 x 3)	Rx64 3.2.3 (raster package) Hijmans and van Ehtten, 2014
Roughness	Difference between the maximum and the minimum value of a cell and its 8 surrounding cells (3 x 3)	" "
Topographic Position Index	Difference between the value of a cell and the mean value of its 8 surrounding cells	" "
Moran's C	Spatial autocorrelation in a neighbourhood (3 x 3)	" "
Textural derivatives (from backscatter GLCMs)		-
Entropy	Measure of spatial disorder in the distribution of elements within the neighbourhood of the Grey Level Co-Occurrence Matrix (all directions)	Rx64 3.2.3 (GLCM package) Zvoleff, 2015
Contrast	Differences of the intensities of the instances within an image in a neighbourhood (all directions)	" "
Dissimilarity	Degree of dissimilarity (Euclidean) in a neighbourhood (all directions)	" "
Variance	Measures the dispersion of the values around the mean in a neighbourhood (all directions)	" "
Statistics (from backscatter)		-
Mean	Average value within a neighbourhood (3 x 3)	base Rx64 3.2.3
Standard deviation	Dispersion of the average value within a neighbourhood (3 x 3)	" "
Minimum	The minimum value in a neighbourhood (3 x 3)	" "
Maximum	The maximum value in a neighbourhood (3 x 3)	" "
Mode	The most frequent value within a neighbourhood (3 x 3)	" "
Median	The median value within a neighbourhood (3 x 3)	" "

3.7.4 Feature selection and model tuning

A feature selection procedure was undertaken to identify the subset of relevant variables from the 14 initial input layers (Table 3). At first, correlation between all pairs of predictors was investigated by computing the Spearman rank correlations between predictor variables, measuring their strength of association. Despite Random Forest being able to handle a large number of highly correlated variables, it has been shown that using only a relevant sub selection of variables improves predictive accuracy as well as computation times (Li et al., 2016). In this regard, a first reduction of the predictor variables dataset (i.e. the information system) was carried out by discarding variables with a correlation coefficient > 95 %. In turn, the Random Forest Boruta wrapper function was used on this reduced dataset (Kursa and Rudnicki, 2010). This function assesses the relative importance of various subsets of input features over multiple runs of the algorithm and provides an estimate of predictor importance (i.e., a Z score). The Z score measures the number of standard deviations a data point is away from the population mean. The significance of the estimate of importance is assessed against the comparison of estimates obtained by original input features and those of the artificial noise features (i.e. “shadow” features produced via permutation of the original variables). This is to overcome the by-chance possibility that a random noise feature could explain part of the variability in the target variable during a single run of the classifier. The feature selection routine was implemented in R (Liaw and Wiener, 2002; R Development Core Team, 2015) using the Boruta built-in functions (Kursa and Rudnicki, 2010).

The RF was applied to both sets of predictors and to predict both classification schemes. Prior to the direct application of the RF classification, a random initialisation partition number (seed) was arbitrarily set to allow reproducibility of results. Using the caret package (Kuhn, 2008), optimal values for the model parameters mtry (the number of variables used at each split) and ntree (the number of trees grown in the forest) were estimated. To achieve this, each model was run multiple times across a range of values for both parameters. The performance of each iteration (considering Accuracy and kappa – see next section) was assessed by a 10-fold cross-validation resulting in 2 splits and 300 trees for the RF_{Simple} model and both classification schemes, and 4 splits and 300 trees for the RF_{++} model.

3.8 Thematic model’s evaluation

Thematic maps do not serve their scope if their information is not directly associated to an objective quantitative measure of accuracy: metrics expressing the “goodness of mapping” allow map’s users and producers to identify the presence, quantity, distribution and nature of the misclassification error, enhancing the utility of the map in a decision-making scenario. Therefore, the accuracy assessment phase of any predictive mapping study should address the following points (Stehman and Czaplewski, 1998):

(1) What is the error frequency: how often does the map not agree with reality?

(2) What is the nature of the errors: which classes are not mapped correctly, and with which other classes are they confused?

(3) What is the magnitude of errors: how serious are they for a decision maker?

(4) What is the source of the errors: why did they occur?

As such, the accuracy of the thematic models produced was assessed in terms of global accuracy (A) with corresponding 95% confidence intervals and Kappa (K) metrics. These indices are derived using the confusion matrix which cross-tabulates observed (ground-truth data points) and predicted (predicted values at the validation sample locations) instances. A confusion matrix reports the correctly classified instances along the diagonal and the confusion between categories in the off-diagonal entries (see Congalton, 1991). Global accuracy is a metric expressing the overall amount of correctly classified pixels, derived by dividing the overall amount of correct allocations by the total number of samples, whereas Kappa measures the difference between the global accuracy of the model and the agreement expected by chance. Kappa ranges between 1 and -1; values close to 0 indicate an inter-rater agreement no better than chance, 1 a perfect agreement and -1 agreement worse than by chance. User and producer accuracies were computed per category. The User accuracy (also referred to as reliability) measures the probability that a prediction of a given category represents that category on the ground (User): this is measured as the fraction of correct allocations and overall number of ground-truth samples for a given class. The Producer accuracy measures the probability of a reference pixel being correctly classified: this is measured as the fraction of correct allocations and all samples predicted as a given class. The No Information Rate (NIR) was also included in the accuracy assessment. This metric relates to the largest category percentage in the data testing against the accuracy achievable by predicting only the majority class. For both modelling approaches, accuracy metrics were assessed against the set of validation points withheld from the overall dataset. Besides the statistical evaluation of accuracy, a visual assessment based on literature and expert and field knowledge was undertaken to investigate further how well the produced thematic models represented reality and better address the previously mentioned points. Visual assessment of the predicted/modelled grid aids the identification of possible errors given that point-wise, the mathematical validation cannot detect errors where control points are not available.

3.9 RESULTS

3.9.1 Data exploration

At first, visual inspection of the representativeness of training and validation ground-truth sample subsets in respect to backscatter and bathymetry is presented (Figure 3.8). The distribution of the sampling effort approximates well the distribution of the backscatter and bathymetry values (Fig. 3.8 A–B) of the entire study area. A slight under sampling of the BS range between -35 and -25 (dB) is visible, as well as a slight over sampling of the lowest BS range around -45 dB. Bathymetry-wise there is an under sampling of the deeper regions (-40 to -30 m LAT) and slight over sampling of the shallower regions (-20 to -5 m LAT). Overall, these observations suggest that the ground-truth datasets represent well the underlying distribution of the MBES primary data. Figure 3.8 C to H shows boxplots of backscatter and bathymetry data grouped by identified substrate classes for three classification schemes: EUNIS Habitat Level III sediment classes (in C and D) and two levels of Folk classes (E and F at 8-classes and G and H at 5-classes after amalgamation). Boxplots were used to assess the class separation potential (discriminative ability) of the backscatter data for both schemes (Figure 3.8 A-F) and to assess the presence of trends. All classification schemes (3, 5 and 8 groups) exhibit a linear increase in median backscatter with increasing percentage of coarse clasts in the substrate classes.
The coarsest class (Coarse in EUNIS habitat level III and sG in Folk) is generally associated with deeper water, while both Folk classification schemes indicate that lesser amounts of gravel (i.e. (g)S, gS, gmS and (g)mS) are also found in the shallow coastal zone. Noticeably, considerable overlap exists between Folk classes gmS, (g)mS, gS and (g)S in Fig. 3.8E. Due to this, amalgamation was done allowing mapping of Folk classes according to the scheme shown in Fig. 3.8G. While the linear trend is maintained, considerable overlap is still present, suggesting that backscatter alone cannot predict the full sediment Folk class spectrum. This lack of discriminative ability is particularly evident for Folk classes with coarser sediments ((g)S-gS) compared to the finer ones (sM-S). The sand (S) class had the largest variability in both backscatter and bathymetry values. Part of this variability, for example in the S class, will inevitably be due to the observer bias and quality of the sample pictures used, but may also be a result of the inclusion of different types of sand in this class (i.e. fine, medium and coarse), as well as biological influence (infauna) and volume inhomogeneities and small-scale morphology below the sonar resolution and left unaccounted by the sample description.

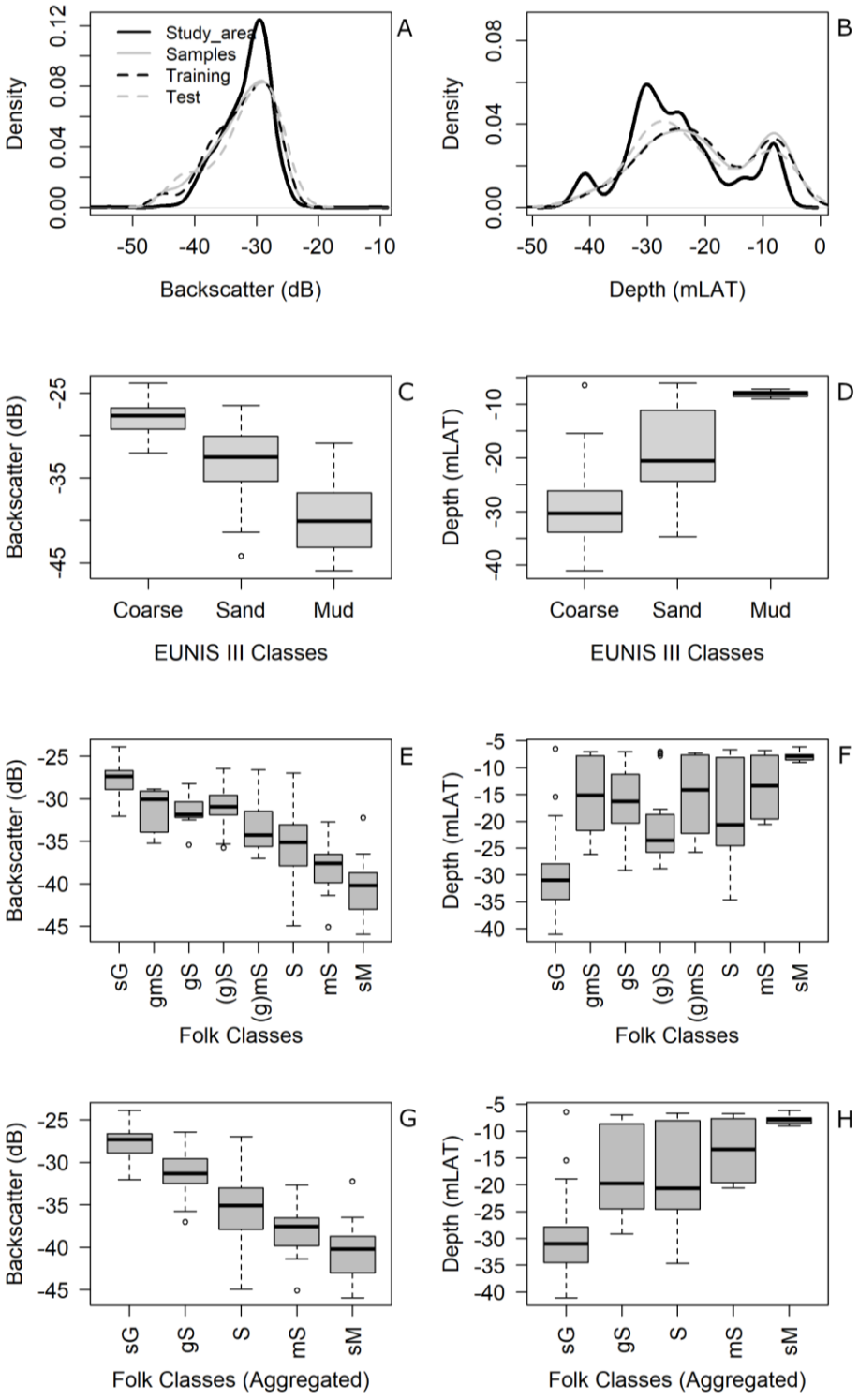


Figure 3.8 – Exploration of overall, training and validation ground-truth datasets. A: Comparison of the backscatter distributions of the study area (solid black line) with that of the entire ground-truth dataset (solid grey line), training (dashed black line) and validation subsets (dashed grey line). B: Same as in A, but for bathymetry. C: comparing the distribution of backscatter values across substrate classes of the EUNIS Habitat Level III classification scheme. D: same, but for bathymetry values. E: as in C, but for an eight Folk class solution. F: as in D. G: as in E for a five Folk class solution, following aggregation; and H: as in F. The black line in the boxplots represents the median value. The upper and lower bound of the box denotes the data that lies between the 25th and 75th percentile. The whiskers denote the full range of the data and outliers are denoted by the black circles.

3.9.2 Relations of sediment variables with backscatter

To gain further insights into the drivers of backscatter intensity in respect to the contributions of the percent weight of fine- to medium sand and coarse sediment fractions, the variability of backscatter intensity was compared with the percentage weight of individual grain-size fractions for the subset of sieved samples (for which the full spectrum of grain-size fractions was considered at the previously defined mesh intervals; Figure 3.9 A-C).

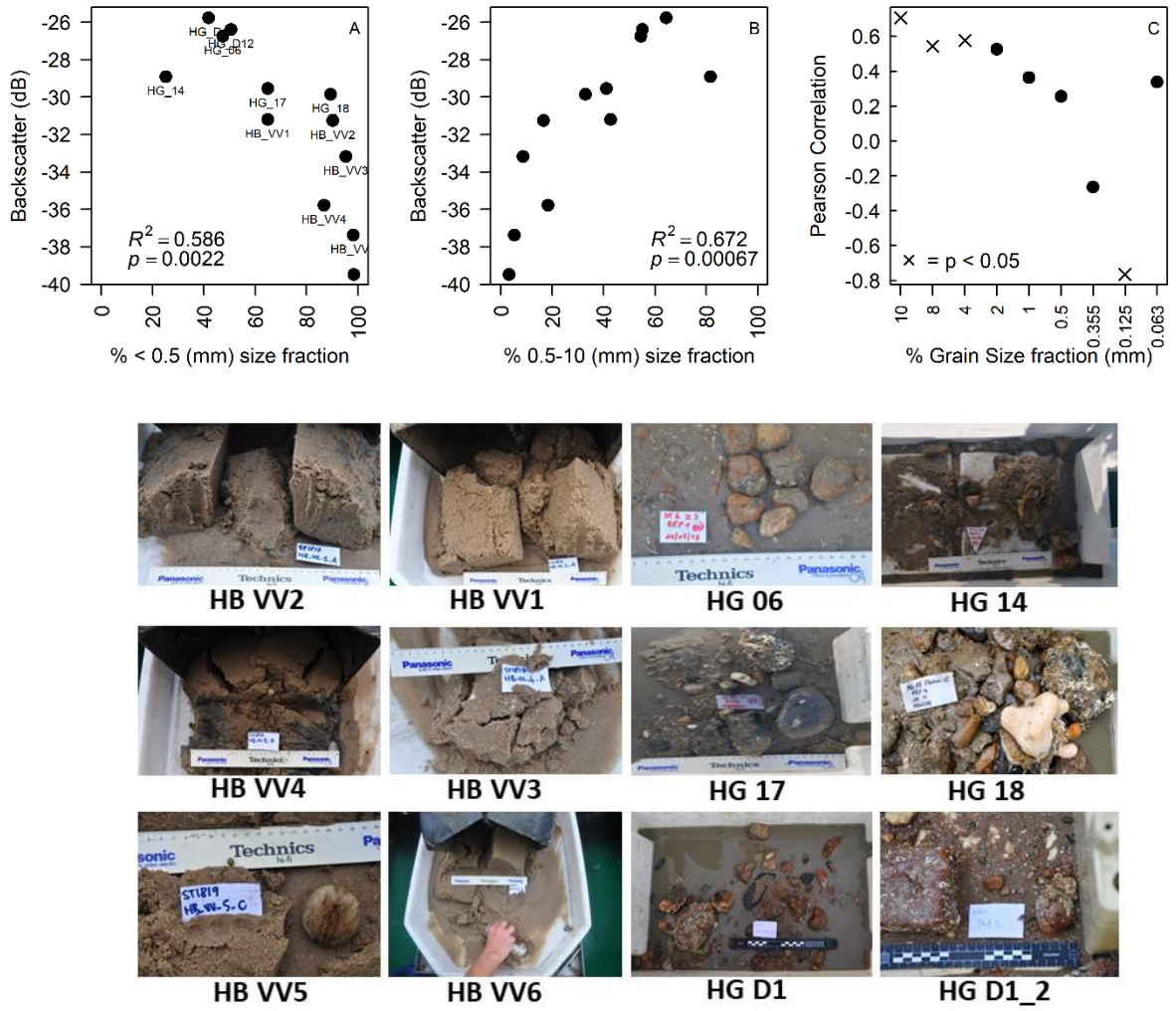


Figure 3.9 – Summary of the relationships between grain size variables and the average backscatter value extracted at the locations of the set of sieved sediment samples (n = 13). A; simple linear regression between the percent weight of the fine sediment fraction (range 0.063-0.5 mm – clay to medium sand) and the backscatter mosaic values extracted at these locations within a 25 m circular buffer; B; same as in A but for the coarser sediment fraction (range 0.5 - 10 mm – coarse sand to gravel) roughly approaching the scale of the wavelength in use (λ @ 300 kHz with sound speed 1500 ms^{-1} = 5 mm). A and B are both significant with R^2 0.58 and 0.67 respectively and p-level < 0.01. C; Pearson correlation coefficient between average backscatter extracted at the sample locations and each individual sediment fraction, reflecting the mesh intervals used in the sediment sieving analysis. Crosses denote Pearson correlations with p-level < 0.05. Corresponding sample pictures are reported below. Sample ID as in A.

In Fig. 3.9-A, the observable trend is a linear decrease in backscatter intensity with increasing weight percentage of the sand fraction in the range of 0.063 – 0.5 mm. In Fig. 3.9-B, the trend observable is a linear increase with increasing percent of the coarse sediment fraction in the range 0.5 – 10 mm. This suggests that increasing percentage of sediment grains approximately around the size of the wavelength (~ 5 mm @300 kHz with a sound speed of 1500 ms^{-1}) leads to positive correlation with backscatter. In Fig. 3.9-C, the Pearson product-moment correlation coefficient assesses the strength and direction of the relationship between mean backscatter and the grain-size classes individually. Again, this suggests that backscatter intensity is significantly positively correlated with the coarse fractions (in the range 0.5-10 mm), shifting towards negative (inverse) correlation with the finer sand grain-size size fraction (in the range ~ 0.5 – 0.125 mm) and reacquiring a positive trend with the finest class fraction (~ 0.063 mm). The latter is unexpected, and it is likely by the fact that this fraction was present only in the gravelliest samples (possibly being biotic matter), thus having a high backscatter response (furthermore it must be noted that the percentage of the 0.063 mm size fraction relates to < 2 % of the overall spectrum: possibly an insufficient number of observations to infer any deduction). Noticeably, in Fig. 3.9-C, the strength of association (or magnitude of increase) between backscatter and % weight of coarse fractions does not increase beyond the 2-mm size fraction forming a “plateau” (a slight increase is noticeable for the 10 mm size fraction). It is worth noting that the upper 10-mm sieve also included larger clasts (i.e. cobbles), considerably larger than the acoustic wavelength. Despite the paucity of observations, it may be surmised that the strength and linearity of this relationship weakens when grain-size fractions larger than the acoustic wavelength dominate the sample. This may be interpreted further by the slight plateau effect in Fig. 3.9B. This observation is corroborated by recent literature findings later in the Discussion. Noticeably, a mere 10-20 % content in coarse sediment fraction increases the backscatter strength considerably (here about 10 dB). This may suggest that for the classes characterised by the presence of gravel, a degree of dispersion/overlap, is expectable given that this sedimentary feature will quite strictly control the acoustic impedance, supporting the previous section in which amalgamation was applied to overcome the non-uniqueness of the acoustic response in respect to Folk type class.

Relationships between backscatter and D50 are displayed in Fig. 3.10. The figure shows a set of simple linear regression exploring the univariate relationships between mean depth and D50 and mean backscatter and D50 by considering the entire sample dataset (i.e. all survey areas; Fig. 3.10 A-B) and a subset of samples with a relatively homogenous water-sediment interface (i.e. predominantly clean muddy to sandy substrate, unimodal and well-sorted, visually scored as “clean/plain”; Fig. 3.10 C-D). Here, the mechanistic trend of sediment distribution in the BPNS becomes apparent with the finer to coarser median sediment diameter following a shallower (nearshore) to deeper (offshore) inverse trend. Positive linear associations are found between backscatter and D50, considerably increasing in significance for

the “homogenous-interface” subset of samples (R^2 = 0.66). Interpretation of Fig. 3.10-B is hindered by the fact that D50 is not representative of the entire samples dataset here since it relates to the sand fraction only. In any case, D50 should be cautiously interpreted against acoustic backscatter since the presence of a single pebble/cobble/boulder within a relatively plain sample, introducing bimodality, could largely affect this parameter. D50 is a strong backscatter predictor only for those samples exhibiting a very homogenous, well-sorted sediment distribution (where all grains are close to the mean) as exemplified in Fig. 3.10-D where a linear relationship between mean backscatter and grain-size D50 is found up until ~ 540 µm.

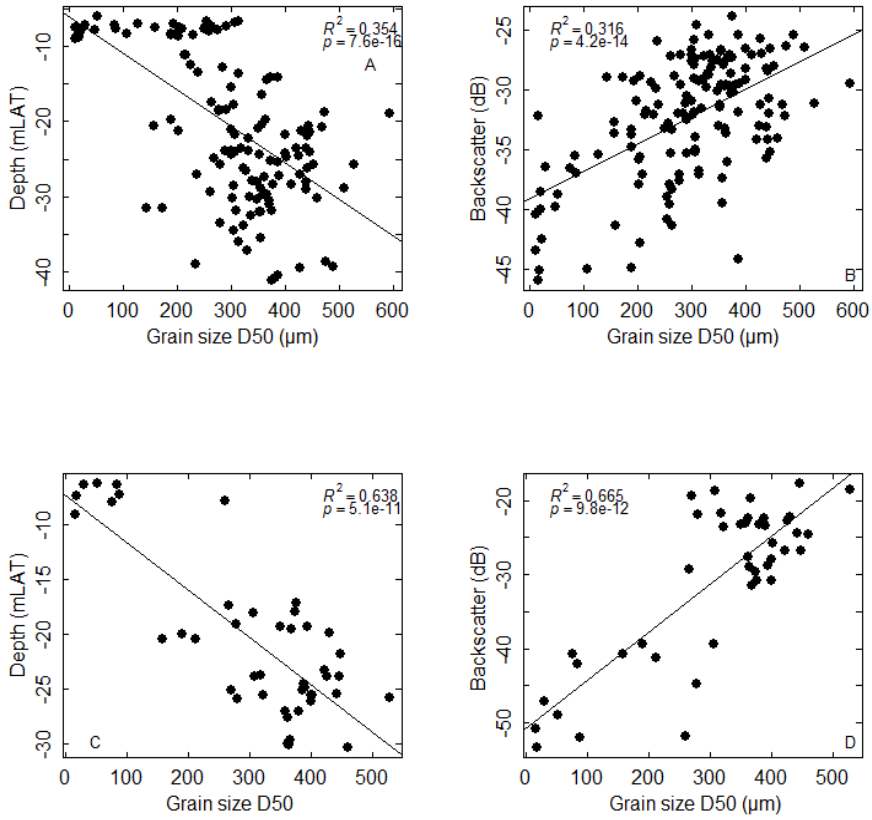


Figure 3.10 – Summary of the relationships between sediment samples and mean depth and mean backscatter at their locations. A and B; simple linear regressions between mean depth and D50 and mean backscatter and D50 respectively for the whole sediment samples dataset. C and D; same as in A and B but for a sub selection of unimodal and well-sorted and visually scored as “clean/plain” sediment samples (i.e. lacking in gravel/shell content).

Together, the findings of the exploratory data analysis provide empirical evidence of the dependency of backscatter intensity and sediment type and agree with a range of studies (e.g. Ferrini and Flood, 2006; De Falco et al., 2010). It becomes clear that within well-sorted, mostly unimodal, fine- to medium-grained sands (< 0.5 mm), the backscatter intensity is dominated by the size of the sand grains, and the D50 parameter is a valuable predictor of substrate type. On the other hand, where samples present coarse/clastic sedimentary features (namely shell and gravel), the backscatter intensity is driven by the weight percentage of the coarse fraction. To

provide further insights of the drivers of seafloor backscatter registered by an MBES, Appendix D presents a multivariate analysis (by means of multiple linear regression) wherein several sediment-related variables are considered together, seeking to understand how the latter may jointly explain the backscatter response.

3.9.3 Feature selection

Figure 3.11 displays the correlation matrix (upper triangle) of the first phase of the feature selection. All backscatter basic statistics (measures of central tendency, minimum and maximum) having correlations > 0.95 % were discarded. Only the Standard Deviation of backscatter was uncorrelated, yet it was discarded in favour of keeping the GLCM Variance, equally measuring the degree of dispersion around the mean value within a kernel. Morphometric derivatives of slope and roughness had a correlation coefficient of 96 %: besides the strong correlation, they were kept due to the differing nature of their morphometric characteristics as well as not being highly correlated to other predictors. The second-phase feature selection routine, Boruta, identified a set of important variables that were kept from the reduced dataset for further modelling using the RF approach. Only variables with a Z score significantly higher than those of permuted variables ($p \leq 0.001$) were kept. For both classification schemes, backscatter was the most important variable identified. Following, in order of importance, variables retained were: Bathymetry, Slope, Roughness and Moran backscatter autocorrelation. The topographic position index, contrast, variance and entropy GLCM textures were found to be irrelevant and were thus discarded from the modelling. The functional relationship of the influence of these variables and the RF_{++Folk} model positive selection of substrate classes are reported in Appendix B where computations of partial dependence plots are shown for each of the five predictors. The partialPlot function from the RandomForest package (Liaw and Wiener, 2002) was used.

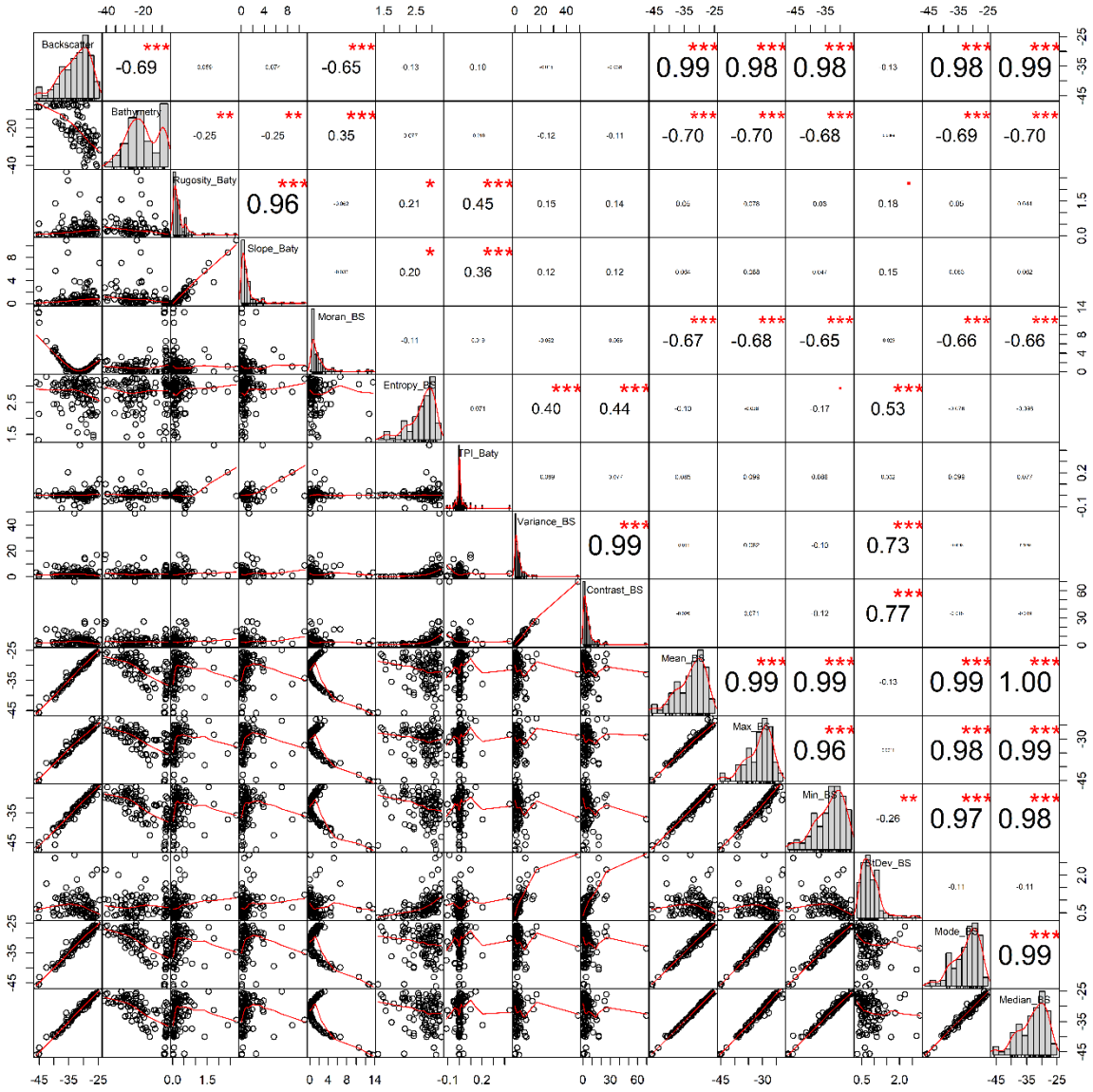


Figure 3.11 – Graphical summary of the first phase of the feature selection analysis. The histogram (grey bars) and its PDF (red line) are shown on the diagonal for each predictor. The lower triangle of the matrix displays the bivariate scatter plots (black points) with a fitted trend line (red line). The upper triangle is the upper triangle of the correlation matrix reporting the correlation values along with the significance levels coded as asterisk and points: "***" = 0, "**" = 0.001, "*" = 0.05, "▪" = 0.1 and " " = 1. The units reported on the X and Y axis should be interpreted as follows: for the top left quadrant (i.e. backscatter histogram with corresponding cumulative distribution), values are reported in the bottom X axis as well as on the right-hand side second Y axis. The values of second variable (i.e. bathymetry), are reported on the top X axis and on the left Y axis. The values of the following variables are equally reported according to the same alternated pattern. The units of the variables can be consulted in Figure 3.7.

3.9.4 Model performance

Metrics of accuracy for each model and classifications scheme are presented in Table 3.3 and displayed in Fig. 3.12 (the corresponding raw confusion matrices are reported in Table A1 along with metrics of by-class model performance (i.e. User accuracy and producer reliability) in Appendix A). These estimates allow to understand not only which classification method performed better in respect to the ground-truth validation set (Overall Accuracy), but also to understand which proportion of the prediction did not occur by chance (k) and to understand the impact of the classes' choice on the accuracy and the way the substrate classes are ultimately thematically represented, giving insight into the trade-off between the accuracy of model predictions and the number and type of features that can be confidently mapped.

Considering the broad-scale mapping of the substrate classes at EUNIS Habitat Level III, the RF_{++} model outperformed the accuracies produced by the k-means and the RF_{simple}, reaching > 80 % in global accuracy and > 70 % in kappa. This is consistent to the aggregated Folk classification scheme where again the RF_{++} model was the most accurate with over 70 % and 60 % global accuracy and kappa statistic. For both classification schemes employed, k-means clustering was the least accurate and produced the lowest values of kappa. The accuracy of this classifier considerably decreased with increasing number of classes (0.56 % to 0.44%).

Table 3.3 - Comparison of model performance on the validation ground-truth dataset. Reported metrics: Global accuracy with 95 % Confidence Intervals, Kappa statistic and No Information Rate.

	Model/Metric	Accuracy	95% CI	Kappa	NIR
EUNIS Habitat Level III substrate classes	k-means	0.56	0.41 – 0.70	0.35	0.43
	RF_{Simple}	0.79	0.65 – 0.89	0.63	0.66
	RF_{++}	0.85	0.72 – 0.93	0.76	0.47

	Model/Metric	Accuracy	95% CI	Kappa	NIR
FOLK Aggregated Classes	k-means	0.44	0.30 – 0.58	0.28	0.29
	RF_{Simple}	0.66	0.52 – 0.78	0.54	0.34
	RF_{++}	0.74	0.61 – 0.85	0.63	0.37

During the modelling phase of this investigation, for the supervised RF approach, it was tested whether omitting a model tuning routine and a per group stratified random selection of validation samples would cause negative effects on the model performance (hence selecting 70 % at random from the entire pool of samples in contrast to stratifying the selection per category – not shown here). In this regard, it was observed that omitting such procedures consistently leads to poorer accuracies

(about 10 % decrease across classification schemes). Furthermore, an RF classification was also attempted for the eight Folk classification scheme resulting in a considerably poorer performance (A = 0.58 [0.44-0.71 95 % CI], K = 0.47, NIR = 0.36 – See Appendix C). This was likely because of two reasons: (1) a lack of sufficient samples to represent each category (an unbalanced sample set), becoming increasingly "rarer" with increasing number of classes. As such, the available ground-truth data failed at supporting the increased complexity of the classification task. (2) Backscatter has a limited discriminative ability of the coarser sediments (as observed in for example Gaida et al. (2018) or Dieising et al. (2014), where the Folk classes are related to acoustic backscatter following class aggregation).

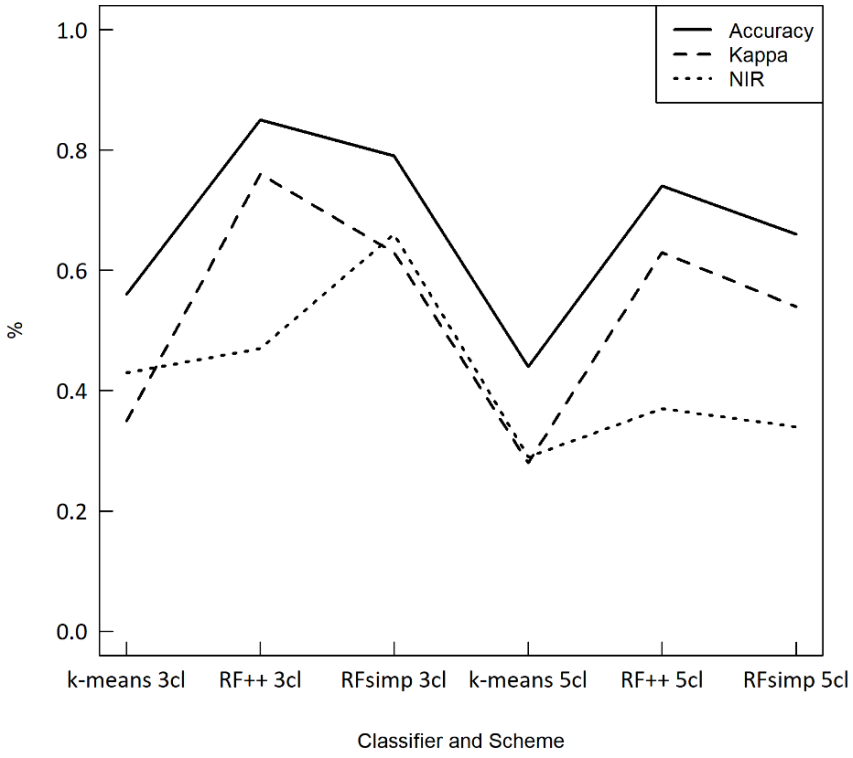


Figure 3.11b – Graphical summary of Accuracy, kappa and No Information Rate for both classifiers and classification schemes.

In the following paragraphs, class specific prediction accuracy refers to the User Accuracy. This is because User Accuracy reflects the percentage of the category that is realistically representing that category on the ground. If only the Producer Accuracy was considered, the misclassified part of the prediction (those areas confused between classes – i.e. in the off-diagonal entries) would not be adequately communicated (User and Producer accuracies are reported in Appendix A along with the raw confusion matrices for each classifier and classification scheme). Class specific prediction accuracies for the RF modelling of the EUNIS Habitat Level III scheme, showed generally a good performance. Sand and Coarse sediments are accurate at ~ 80 and 90 % for the RF_{++} although the coarse class decreases

noticeably in the RF_{simple} model. The Mud class has a lower accuracy in the RF_{simple} model (0.57), though it increases in the RF_{++} prediction (85 %). At this hierarchical level, the classification by k-means produces similar results except for the Sand class predicted at 34 %. The "Sand" category has the largest variability in both backscatter and bathymetry values and it is in fact the most frequently confused category (with Mud and Coarse categories) by the three modelling approaches. Noticeably, at the EUNIS level III classification using the RF modelling approach, the NIR considerably improves by including the set of identified relevant layers. This implies that the model improves in what it is less skewed toward accurately predicting only the most common class (Sand), and in fact, the predictive accuracy is better distributed across classes in the RF_{++} result. Looking at the finer hierarchical level, the modified Folk classification scheme, RF modelling approaches are again producing the most accurate predictions for the five identified classes. Classes sG and gS are the most accurately predicted at ~ 80 % by both RF models. Classes mS, S and sM produce weaker accuracies in the RF_{simple} model and they increase substantially for the RF_{++} model, particularly for the sM class which doubles. The Sand class also improves while the mS class prediction remains stable. The k-means approach at this hierarchical level produces the lowest accuracies, failing at predicting entirely the S class. Overall, for both classification schemes, the RF_{++} produces the highest global accuracy, kappa statistic and per-class user accuracies.

Figures 3.12 and 3.13 illustrate the classification results for the EUNIS Habitat Level III (Fig. 3.12) and modified Folk (Fig. 3.13) classification schemes employed for the k-means (A series), RF_{simple} (B series) and RF_{++} (C series) models and for three selected areas. For both figures, the A group displays the Hinder bank study site; A1 group the Ostend disposal ground of dredged material, and the A2 group the Thornton bank study site (Figure 3.1B shall be used as a reference for the geographic locations of these three areas within the BPNS). A visual inspection of these maps, draped by hill-shade layers, aids the interpretation of the quality of the predictions not accounted for by the statistical accuracy assessment (i.e. due to a lack of datapoints and to interpret the classified sediment in respect to the geomorphology) and to identify the most prominent differences between classifiers and classifications.

The most striking observation is that across both classification hierarchies (EUNIS Habitat Level III to modified Folk) and for k-means and RF_{simple} classifiers, the top of the sand dunes (within the Hinder bank region; Oost-, West- and Noord Hinder) are consistently mapped as Mud and/or sM (i.e. predominantly muddy sediments) depending on the scheme employed. In the BPNS muddy sediments are predominantly distributed in the coastal area (South East) where the water turbidity maximum zone occurs (Fettweis et al., 2006). A paucity of samples in this area leads to the statistical assessment not fully capturing this important classification error. A visual assessment aids the interpretation of this error. This misclassification could result from the similar acoustic backscatter mosaic values found in the predominantly

muddy areas (i.e. Ostend study area in the A1, B1 and C1 series of Figures 3.12 and 3.13) and that found on top of the sand dunes in the Hinder bank region. As observed in Montereale-Gavazzi et al. (2017), where the Hinder bank region was the focus of a change detection study, the top of the sandbank is the most dynamic part of this morphological feature. Sediments are composed of very-well sorted, highly mobile fine to medium sand and likely have a considerably higher water content (water-saturated) than the adjacent flank and swale portions of the study site (where roughness and surface scatterers are the dominant scattering mechanisms). From an acoustic perspective, this would result in relatively lower acoustic impedance contrasts between the water-sediment interface and the water medium, leading to similar in-sediment absorption contributions as those encounterable in the unconsolidated and coastal muddy areas. MBES acquired over the sand dune systems of the North Hinder study site were logged during high tide (otherwise being hazardous for navigation at low tides) and are thereby representative of the most dynamic moment of the tidal phase when considerable sand transport and mobility are expected, explaining the shared similarities in the water-sediment interface status between coastal muddy and the offshore top zone of sand banks and sand dunes (this observation is confirmed by similar classification studies of the Belgian sandbanks substrate type [Roche, pers. Comm.]). The RF_{++} model almost entirely overcomes this misclassification at both hierarchical levels mapped as the class separation potential was substantially increased by including a set of relevant predictors which enhanced the array of rule-based decisions of the classifier. Noticeably, while the deeper sand dunes (i.e. SE and NW, having shallowest depths of -12 and -14 m LAT, respectively) are correctly classified by the RF_{++} model, a slight misclassification is still present in the central portion of the study area, on top of the shallowest (-7 m at high tide) sand dune complex. Clearly, the shallowness of this morphology under the influence of peak current velocities at high water (reflecting survey times), makes it the most dynamic part of the study area, subject to considerable sediment mobility and water saturation.

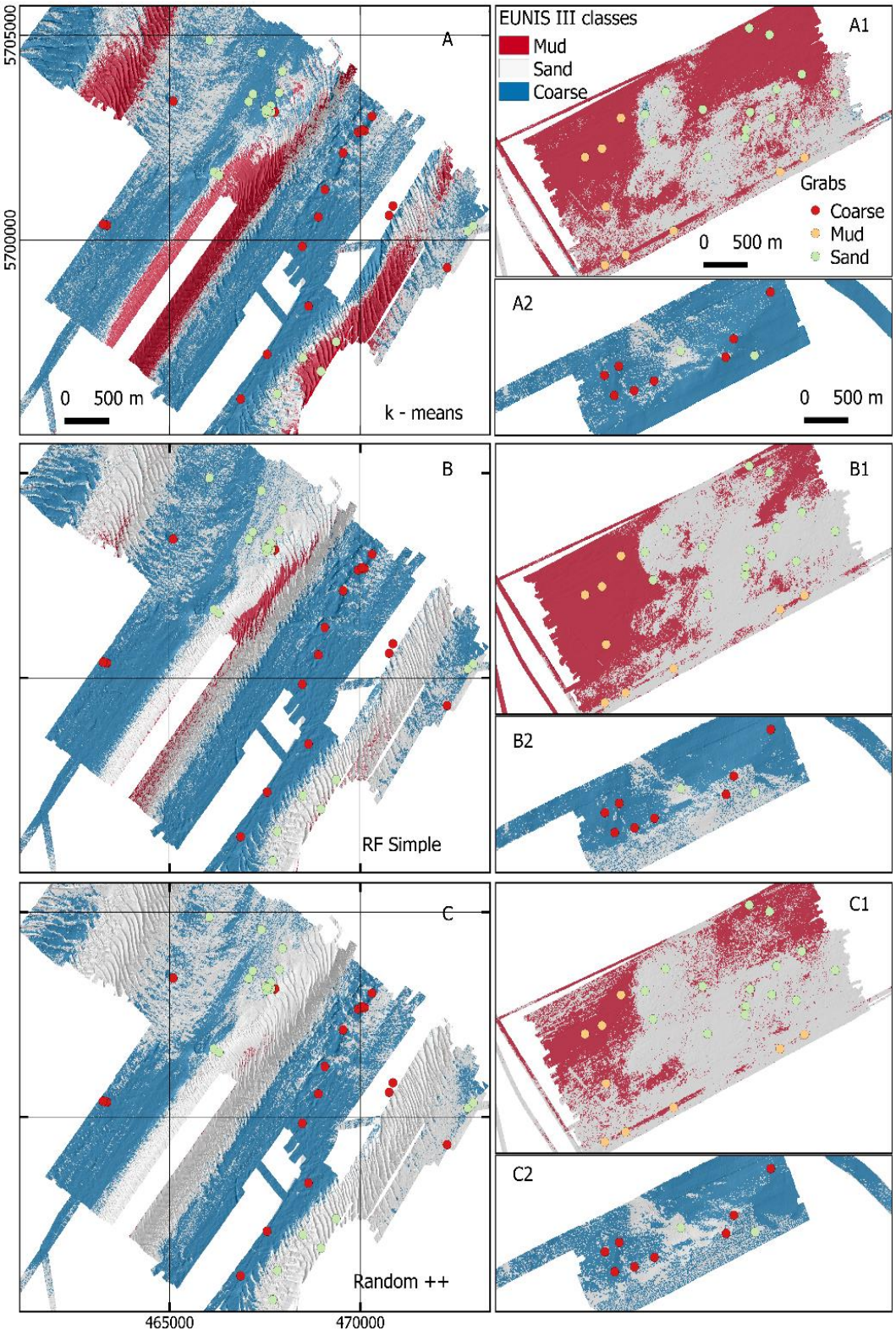


Figure 3.12 – Maps produced by k-means and RF models for the EUNIS Habitat Level III classification scheme. In the first column; maps produced by A: k- means, B: RF $_{simple}$ and C: RF$_{++}$ within the offshore Hinder bank region. In the second column; same models with details of the Ostend disposal ground of dredged material and the Thornton bank area survey areas. A hill shade layer derived from the bathymetry is superimposed to capture the distribution of sediment in respect to the underlying morphology.

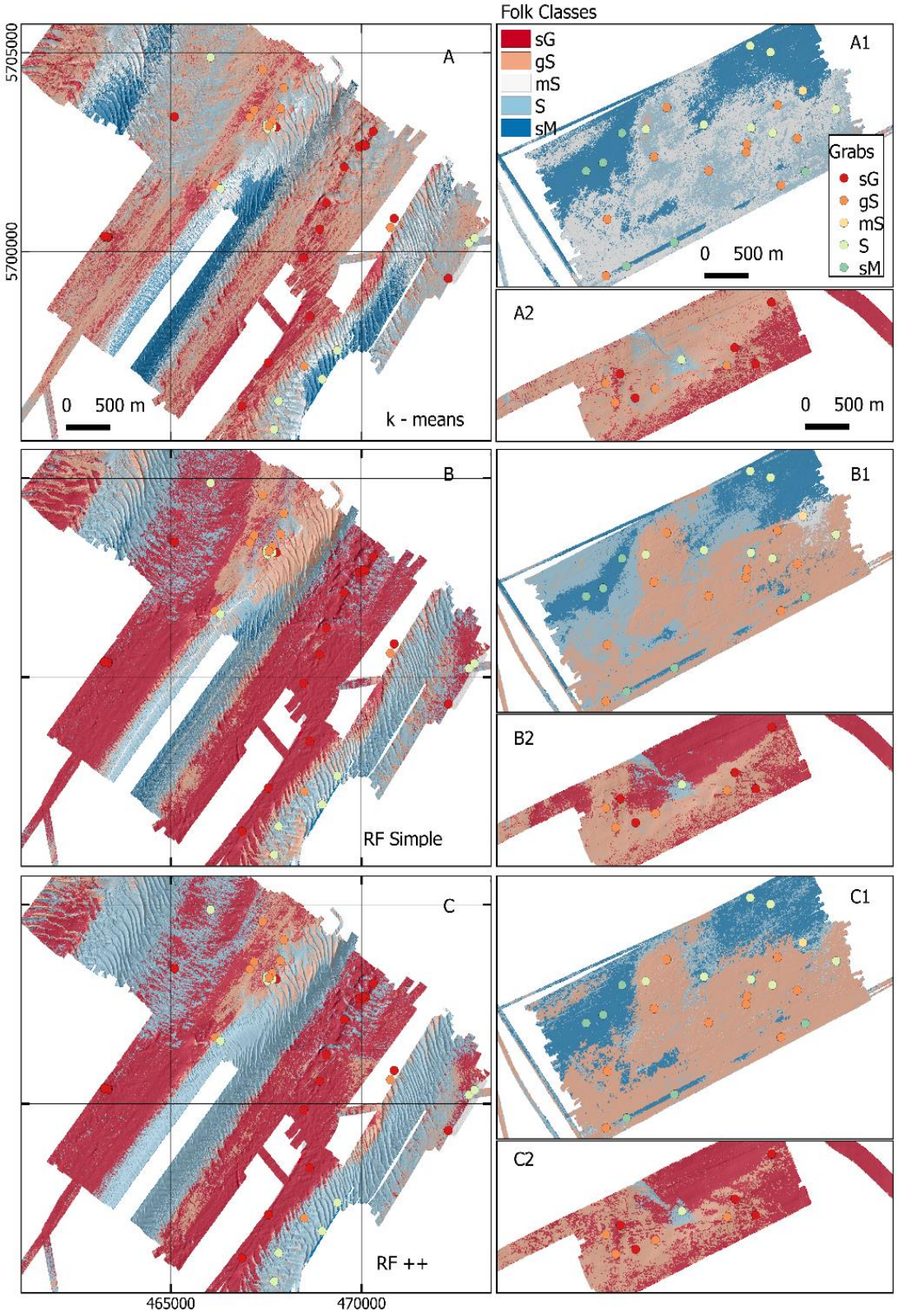


Figure 3.13 - Maps produced by k-means and RF models for the modified Folk classification scheme. In the first column; maps produced by A: k-means, B: RF_{simple} and C: RF_{++} within the offshore Hinder bank region. In the second column; same models with details of the Ostend disposal ground of dredged material and the Thornton bank area survey areas. A hill shade layer derived from the bathymetry is superimposed to capture the distribution of sediment in respect to the underlying morphology.

A last and important observation regarding a further potential source of thematic error (i.e. misclassification) regards the orientation in which the backscatter data have been acquired in respect to the underlying morphology: the azimuthal dependence on backscatter. Depending on the isotropy or anisotropy of the seafloor morphology, the survey azimuthal direction is a factor known to largely affect the backscatter strength (Ferrini and Flood, 2006; Lurton et al., 2018). Noticeably, the presence of sand ripples beyond the imaging resolution of the sonar (i.e. tide-driven micro-oscillatory ripples), will cause a different acoustic response when surveyed either with the ship's heading parallel (increasing the backscatter strength as a direct consequence of the ripple flanks being normal to the across track beams) or perpendicular to the ripple's crests (including at varying azimuthal intervals e.g. increments of 15°). Dedicated experiments demonstrated how this effect can alter the strength of the response by up to 12 dB in the incidence angle range ~ 20° to 40° (Lurton et al., 2018). It is however important to keep in mind that where the angle dependence has been compensated and referenced to 45° incidence, as for the backscatter mosaic production phase in this investigation, the impact of the small-scale ripples will be cancelled out and the backscatter level will not depend on the survey heading. In other words, the same seafloor type will provide the same average backscatter response irrespective of the presence and/or absence of ripples. This should be perceived as an advantage where the classification target is the sediment type, though it may be a limitation where the classification target is the interface status (i.e. to whether rippled or not). Presence of such fine scale structures remain beyond the imaging capability of the sensor, leaving analysis of the angular response as the sole means of detection.

3.9.5 MBES substrate maps

Hereafter, a description of the observed patterns in sediment distribution for the best performing substrate predictive model of the modified Folk scheme (RF_{++Folk}) is presented. The consistency of the classification is compared to previous (i.e. older generation echosounders, processing and classification routines) mapping research in the BPNS. A set of 10-m horizontal resolution ASC maps made available by the Federal Public Service Economy, Continental Shelf Department (CSD) of Belgium (Koen Degrendele and Dr. Marc Roche) and based on previous generation Kongsberg EM1002 backscatter data are used for comparative purposes. These maps were classified with the Kongsberg Triton module used at the time. Classes were identified by the combined interaction of quantile, pace and contrast indices which were used to define clusters (in Roche, 2002 and Bellec et al., 2010) using the Kongsberg Triton supervised classification module (Kongsberg Simrad, 1999-2001). The comparison shall enhance the appreciation of the spatially explicit character and accuracy of the models achieved in this investigation in respect to former generation approaches. Figure 3.14 displays details of backscatter and bathymetry signatures of the five main categories identified in the RF_{++Folk} approach and mapped across the entire extent of the seamless MBES dataset. The predicted percentage cover of

each class accounts for: sG = 33.3 %, gS = 27.3 %, mS = 1.16 %, S = 35.9 % and sM = 2.1 %. Muddy sand (mS) and sandy Mud (sM) classes are the least represented (overall 3.26 % of the coverage) given the surveys planning was predominantly "skewed" towards mapping areas of "coarse" substrata (gS-sG).

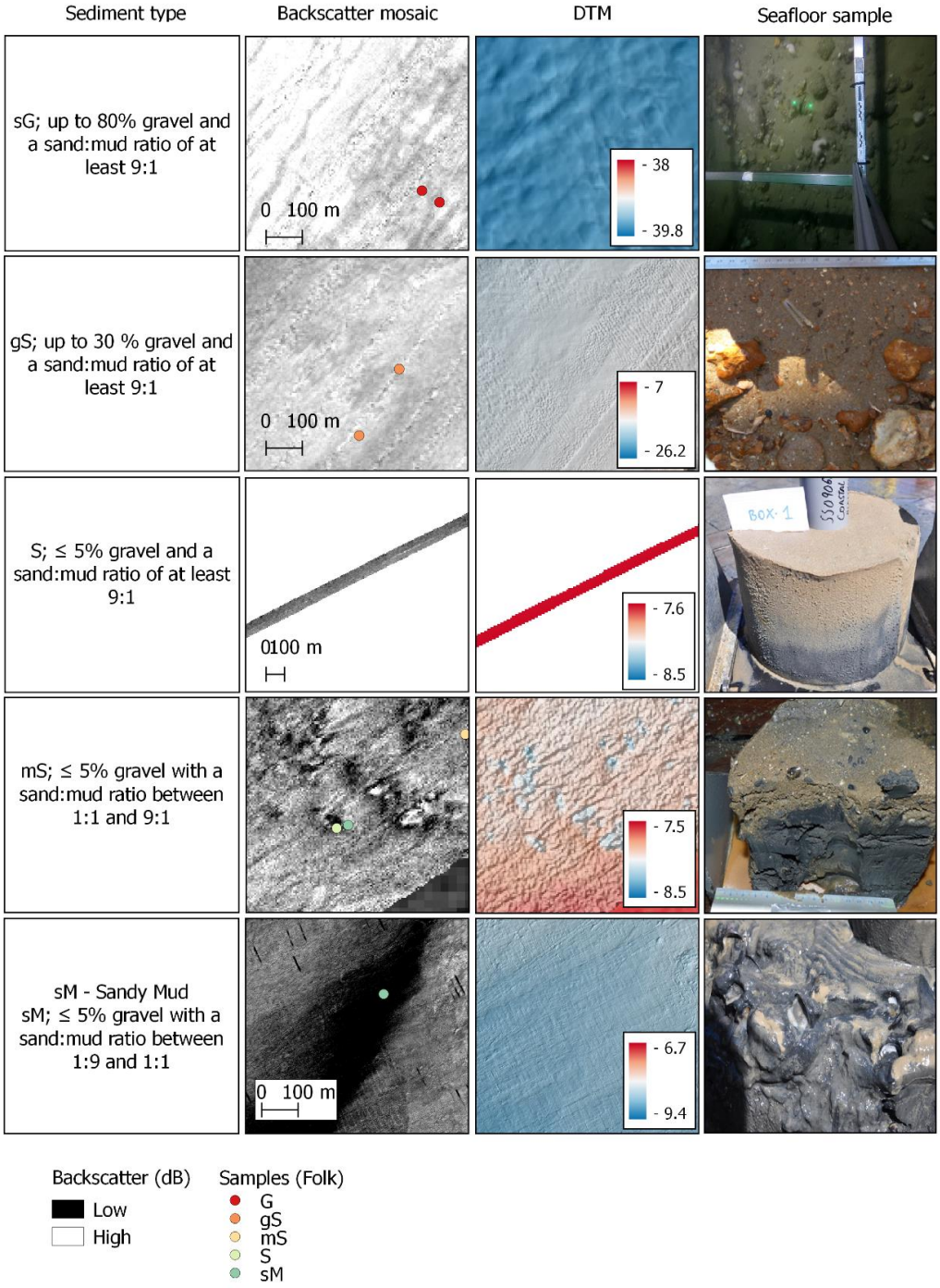


Figure 3.14 - Distinctive backscatter and bathymetric signatures present within the study area and correspondent ground-truth images and locations. In the top right quadrant (class sG), the laser beams in the video image are 10 cm apart. The ground-truthing and backscatter mosaic details should be interpreted as "generally" describing the acoustic class. Variability is to be expected since certain classes were amalgamated (e.g., classes gS and sG).

Although the sG class had the highest individual accuracy, it represents a mixture of various typologies of gravel areas, including areas predominantly characterised by a dense cover of coarse bioclastic detritus (namely shell fragments/hash), as well as areas of dense clusters of pebbles and cobbles ("pebble and boulder fields" as identified in Gilson's (1907) pioneering studies on the Hinder banks) with varying sand and/or clay enrichment (this variability is clearly represented in the ground-truth pictures reported in Figure 3.15). Both gS and sG classes should be interpreted as highly variable given the limitations identified in the exploratory data analysis. As identified in the exploratory data analysis, from an ASC perspective, such relatively subtle differences cannot be distinguished in the sG class as the high percentage of coarse fraction proportion will predominantly drive the acoustic response in the backscatter, masking the fine sediment fraction contribution irrespective of its nature (i.e. presence of mud and/or sand will not be captured by the acoustic data and the backscattering strength will be primarily driven by the coarse component). Furthermore, thin (order of centimetre) sand veneers/patinas covering gravel might also result in the classification as sG given the strong volume echo contribution of the buried subsurface (as also observed in Todd, 2005). This could occur in the North hinder region (Figure 3.16) where the sedimentary Holocene cover interpreted from seismic surveys (hence at a metric vertical resolution) is expected to be < 2.5 m (Deleu and Van Lancker, 2007) including areas with presumably thinner thickness (Roche, 2002; Personal communication based on sampling campaigns). Nevertheless, for sandy substrates, it is expected that the penetration into the sediment for a 300 kHz operating frequency (as the data herein used) is limited to ~ 3 cm (Huff et al., 2008). Similarly, the gS class is highly variable and represents areas where the coarse sediment fraction varies between ~ 5 and 30 %. The degree of generalisation of these patterns has to be interpreted in respect to the scale at which they are represented i.e. 25 m^2 patches.

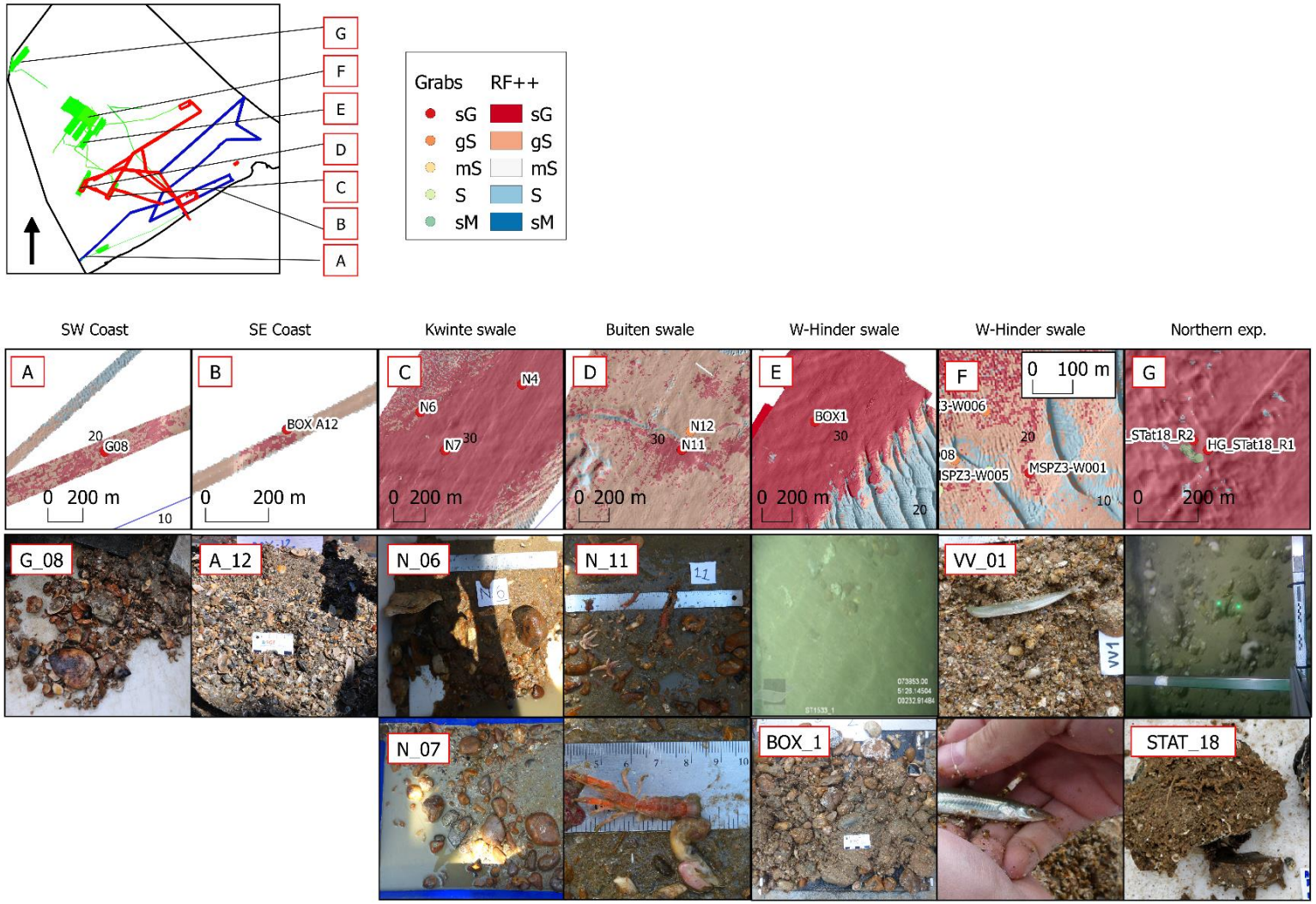


Figure 3.15 – Details of areas predicted as sG in the RF_{++Folk} model. Geographical reference to each quadrant is given in the top left map of the BPNS. Sample pictures correspond to ground-truth locations on the map for which the label is displayed on both the map and picture. Identifiable macrobenthic species in D: Upogebia deltaura, Asteroidea; F) Ammodytes sp.; G) Alcyonium digitatum, Sabellaria sp.

3.9.5.1 Hinder Banks regions

Figure 3.16 shows the resulting classification for the surveys carried out within the Hinder bank region. As visible by the ground-truth pictures and hill shade overlay, the category Sand (S) is correctly allocated to the sand dunes of the Oost- West- and Noordhinder system. The occurrence of gravel is mostly restricted in the swales where a typical “hillocky” terrain morphology is visible, particularly in the southern part of this study site where considerably higher amounts of gravel are known to occur (Roche, 2002; Van Lancker et al., 2007). Towards the North, after the East-West traversing scarp (yellow dashed line in Fig. 3.16 dividing the study area), the seafloor becomes increasingly sandier and the gravel is mostly restricted within the throughs between sand dunes (in the NW). The hillocky character of the terrain is mostly present in the southern part of the study area where, at the foot of the Oost hinder, a system of barchanoid dunes has evolved. The latter morphologies are generally associated with the presence of coarser sediments (Todd, 2005) resulting, for this region, in relatively dynamic features with migration rates of ~50 m over ~10 years (Montereale-Gavazzi et al., 2017). Houziaux et al. (2008) referred to the troughs in between barchan dunes as “gravel refugia of biodiversity”. Since they occur near the lee side of the steep barchan dunes, bottom trawling fishing gears are hypothesised to merely jump over the dunes, hence without bottom impact. Due to this, it is found that the majority of epifaunal assemblages colonising exposed gravel are distributed in these troughs. Noticeably, and supporting this observation further, column F in Fig. 3.15 shows a flat area of seafloor situated between sand dunes (situated at the NE base of the central part of the West-hinder complex) and characterised by a speckled pattern of alternating sG and gS substrate classes. This speckled pattern corresponds to sand eel (Ammodytes sp.) habitat (visible in the bottom cell of this column). Numerous specimens were sampled in the Van Veen grabs acquired within this area. Sand eels are important fish species, sought after by the fishing industry and an important foraging source for several fish and bird species up the food-web (van der Kooij et al., 2004).

3.9.5.2 Flemish Bank region

Despite the CSD classifications being largely affected by striping artefacts, owing to the former generation of backscatter angular compensations applied in the Kongsberg Triton module (recently overcome by the introduction of better compensation algorithms (see Fonseca and Calder, 2005)), the agreement between classifications is qualitatively good with both instances capturing the progressive distribution of sG and gS within the innermost part of the swales in between the Kwinte and the Buiten Ratel and the Buiten Ratel and the Oost Dijck bank (Figures 3.17 and 3.18). Noticeably, transit parts of the RF_{++Folk} surveys correctly identify the S class, crossing areas of sand dunes (as visible from the hill-shade overlay).

3.9.5.3 Thornton Bank region

Further visual agreement can be gleaned from the Thornton bank swale area (in between Thornton and Goote bank; Figure 3.17). Here, overlay of the RF_{++Folk} model shows good agreement for the sG class and the transition from sG to gS and S is noticeable in the NE part of the map (in the trajectory survey for the 17-322 campaign using the EM2040D). Coincident patches of sG class are also visible between classifications in, for example, the western zoomed-in quadrant (black outlined quadrant; Fig. 3.18) as well as in the Eastern quadrant capturing a sand pocket. Noticeably, along the ST1817 EM3002D trajectory (NE to SW; Black contoured polygon) the RF_{++Folk} gS class (in orange), follows a very similar pattern to that of the gS and (g)mS classes (light grey and pale violet, respectively) in the former Triton classification.

.

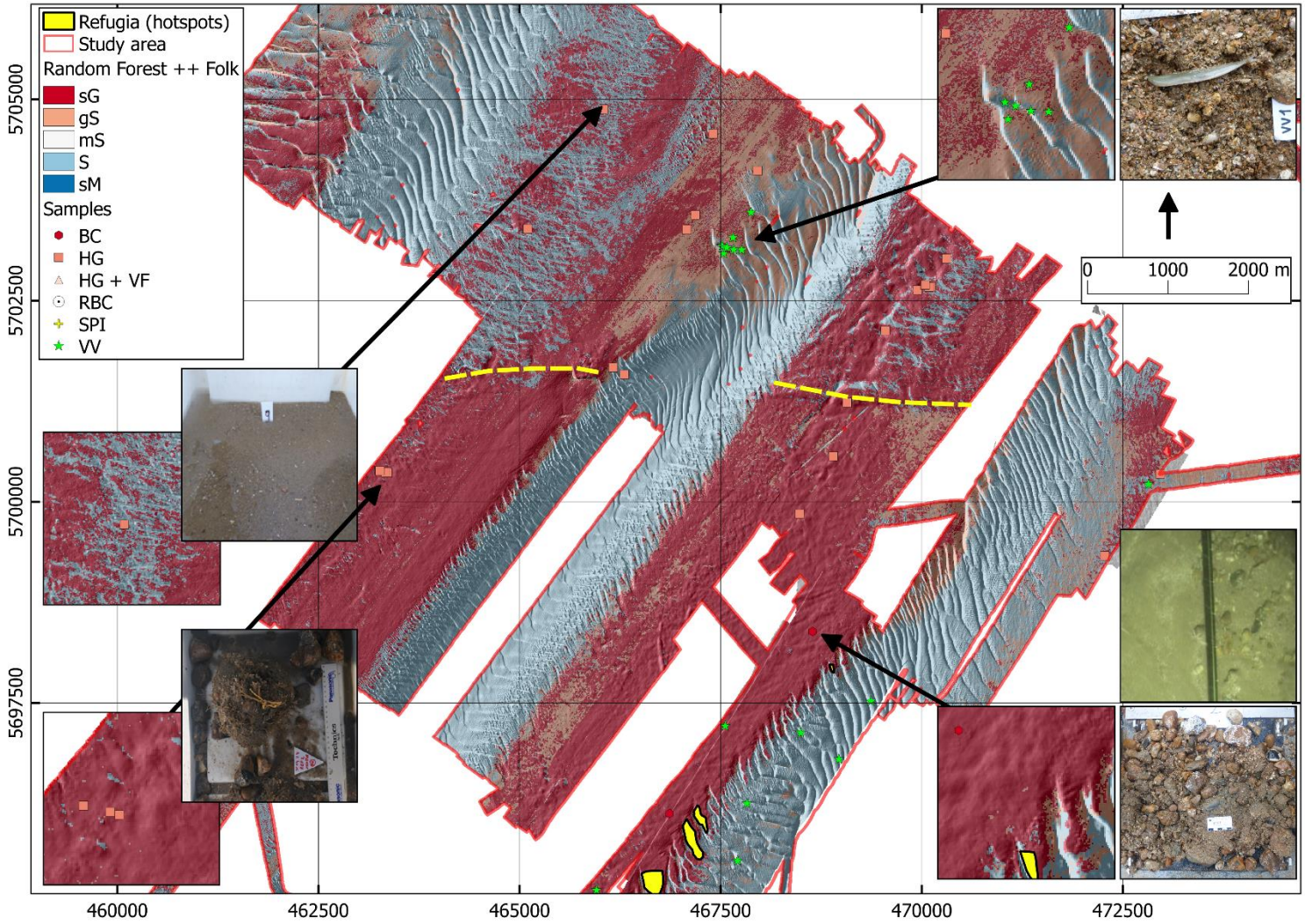


Figure 3.16 – Seafloor substrate classification of the Hinder bank region. Details of predictions along with ground-truth pictures are displayed. The yellow dashed line denotes the location of the scarp delineating a former paleovalley of the Rhine-Meuse river system (De Clercq et al., 2016). The yellow polygons denote the location of “refugia” hard substrate epibenthos hotspots identified during previous oceanographic campaigns (Van Lancker et al., 2016).

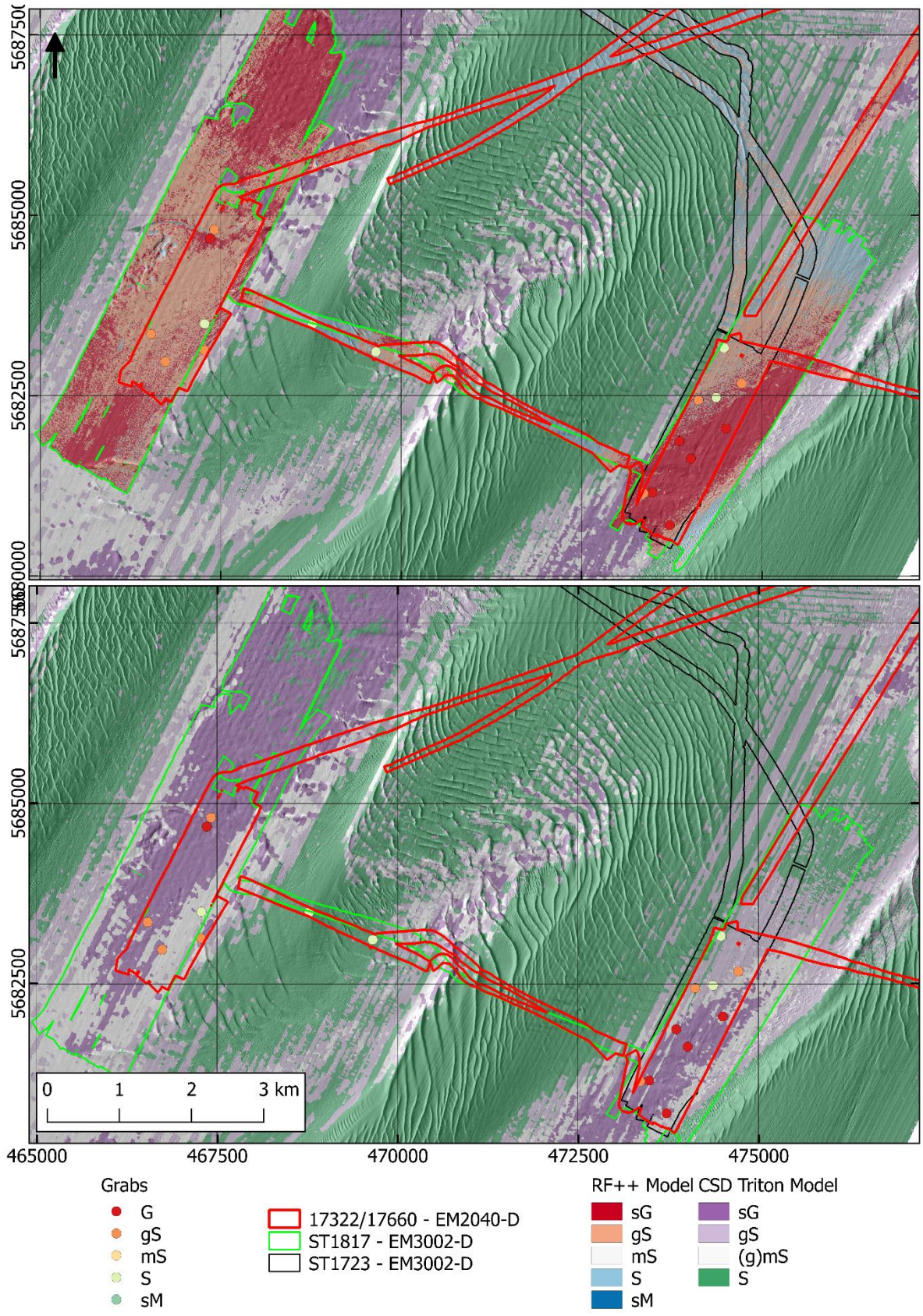


Figure 3.17 – Seafloor substrate classification of the swales in between the Kwinte and Buiten Ratel banks and Buiten Ratel and Oost Dijck banks. RF_{++Folk} predictions (upper quadrant) and CSD classifications in Roche (2002) (lower quadrant). A hill-shade layer provides a perspective view of the morphology. Background bathymetry courtesy of the CSD.

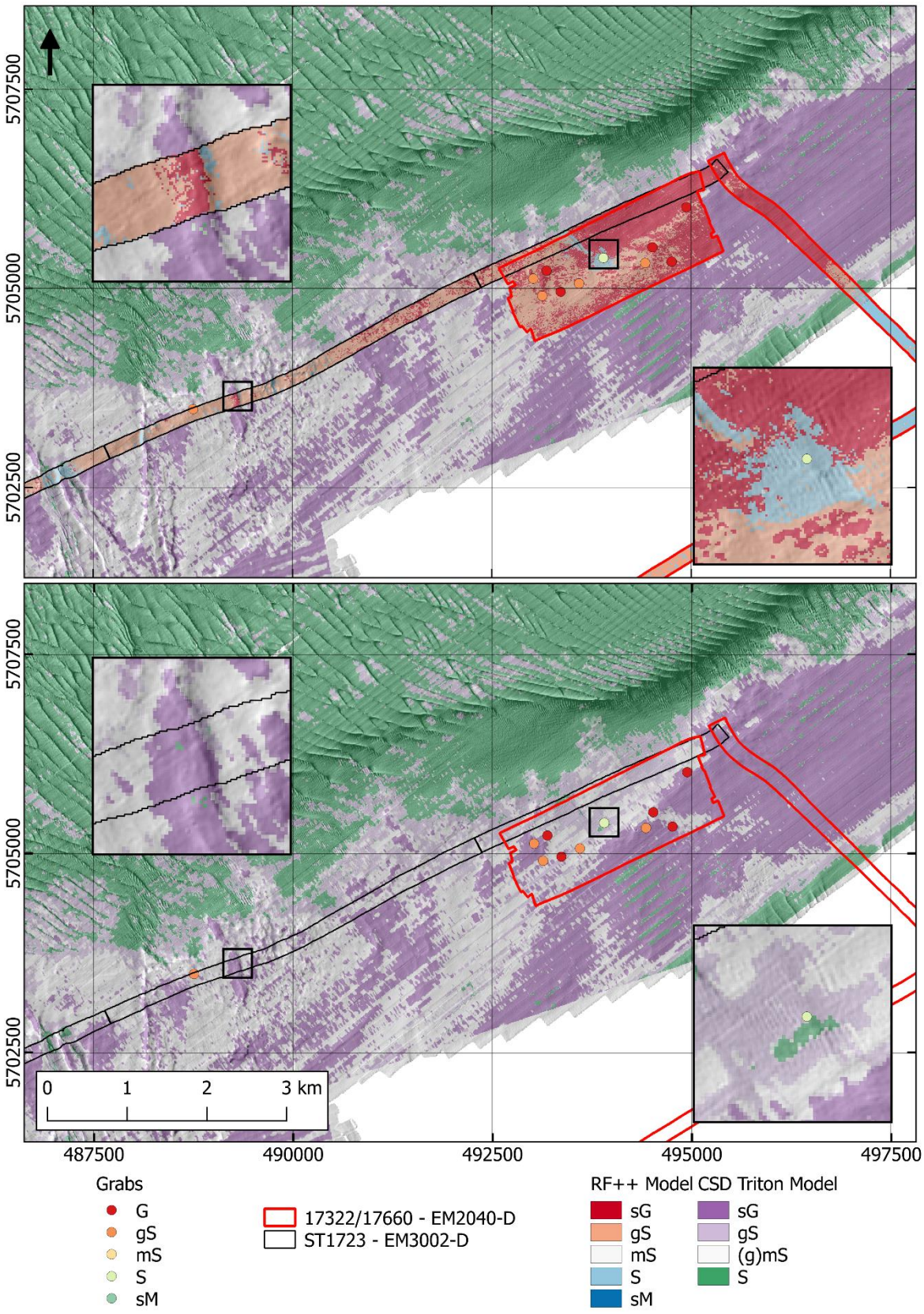


Figure 3.18 – Seafloor substrate classification of the swale in between Thornton and Goote banks. RF_{++Folk} predictions (upper quadrant) and CSD classifications in Roche (2002) (lower quadrant). A hill shade layer provides a perspective view of the morphology. Background bathymetry courtesy of the CSD.

3.10 Interpretation of the modelled angular responses

Backscatter average angular profiles were generated for each sampling location by implementing the methodology described in the Methods section after e.g. Lamarche et al. (2011) and Fezzani and Berger (2018) and subsequently grouped by the five Folk classes identified. Figure 3.19 displays a set of selected angular profiles for the Folk-5 categories. These samples/locations were selected based on the across-track homogeneity of their corresponding compensated backscatter imagery. Noticeably, despite the relative comparability of the backscatter values between echosounder systems (those acquired by the EM3002D are fully comparable given the thoroughly assessed system stability), a low to high (BS) gradient is established from sandy mud (sM) to sandy gravel (sG) classes (note how in the oblique angular range the backscatter increases with increasing grain size). The five classes have distinct angular profiles exhibiting distinct shapes and values. This confirms the physical and sedimentological differences between the classes identified in the ground-truth data and relate to differences in grain size, volume heterogeneity and roughness (as apparent from the corresponding ground-truth pictures). The angular profile belonging to the sG (sandy gravel EM3002D – acquired in the Northern exploration area; see Fig. 3.1a for survey areas locations) sample exhibits a lack of specular amplitude and a wide and very high Lambert-like distribution relatively homogenous across the swath, irrespective of the angle of incidence. This kind of across-swath homogeneously distributed and high backscatter level is indeed expected given the rough and coarse nature of the sediment type at this sample location (note the ground-truth picture in Fig. 3.19), mainly composed of gravel (pebbles to boulders) with evidence of calcareous bio-encrustations by serpulidae polychaetes. The next class, gS (gravelly sand, EM3002D – acquired in the Westdiep area), shows a similar angular profile to that of the sG class though has a lower backscatter strength across the swath as well as exhibiting a slight specular amplitude. Indeed, this reflects the nature of the sediment type, here characterised by abundant shell fragments (bioclastic detritus) nested in a predominantly sandy sediment matrix: relatively smoother and flatter than the angular response presented for the sG class. The following angular profile is that of the S class (Sand – EM3002D – acquired along the MSFD coastal trajectory). It has lower BS values across the swath, a more pronounced specular reflection and extent, denoting the smoothness of this area, as well as a faster BS diminution over the oblique and fall-off grazing angles. The class mS (muddy sand – EM2040-S - acquired in the MSFD coastal trajectory) and the class sM (sandy mud – EM2040D -acquired in the MOW 1 pile area) show typical angular profiles for muddy and fluid-like sediments with a smooth and flat distribution (the grain size of these areas cannot support a significant roughness), showing a very strong specular peak and a very fast decrement of the oblique and fall-off grazing angles. Indeed, this is apparent in their corresponding ground-truth sample pictures, showing a fine (< 1 cm) sand veneer on compact mud (mS class) and an unconsolidated, fluid-like muddy (sludge-like mud) sediment for the sM class.

While this ad-hoc selection of homogenous sample locations shows the clear hydro-acoustic, physical differences between sediment types, high within-class variability does occur in this study (see Figure 3.21). This is as also observed in other studies, such as that of Fezzani and Berger (2018) where single sediment classes exhibited a variability range of up to 10 dB at 45°. This points to the fact that complexity is indeed the rule rather than the exception in marine sediments and that while a five-class scheme could be accurately predicted using the RF machine learning approach, describing the general distribution of five commonly used seafloor sediment categories, increasing the sampling effort would inevitably increase the complexity of the map produced, possibly allowing an increased discrimination and thus the prediction of a higher number of classes (i.e. the sub-groupings inherently present within a sediment domain). This would be of great interest in view of detecting changes of interest (i.e. due to the impact of an anthropic activity).

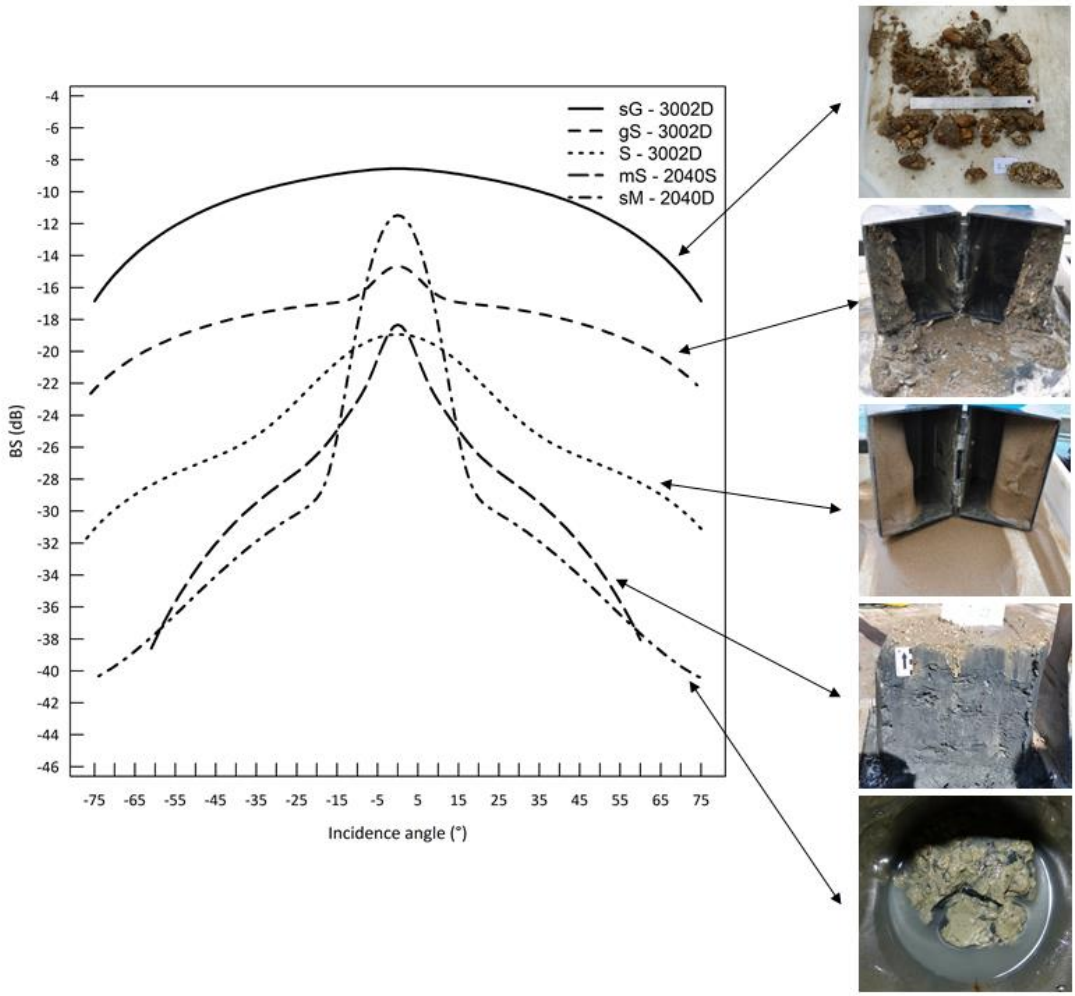


Figure 3.19 – Backscatter strength angular profiles fitted with the GSAB model for five selected samples for each Folk category (Folk 5 scheme). The corresponding ground-truth images are displayed aside. Note that the comparison between curves (except for those from the EM3002D system) are only relatively comparable since no proper intercalibration of the three systems was carried out. Also note that the EM2040-Single head system swath covers only up to 60° port and starboard, as oppositely to the EM3002D and EM2040D extending up to 75°.

Figure 3.20 shows the average modelled angular profiles for each category (again referring to the Folk-5 scheme) and summarised by echosounder model. The good class separation is evident for the EM3002D system with a > 2 dB separation between curves at 45° and a progressive widening of the specular amplitude and decrease in lambert-like behaviour from the coarser to the finer class (only three classes were available within this dataset). For the EM2040-S and Dual systems, the good class separation is still apparent with a similar behaviour, although strong overlap is observable. In the EM2040S, classes S and mS are very similar both in shape and backscatter levels whereas for the EM2040D, classes mS and sM are practically identical. Potential explanatory factors/limitations are discussed in the forthcoming Discussion.

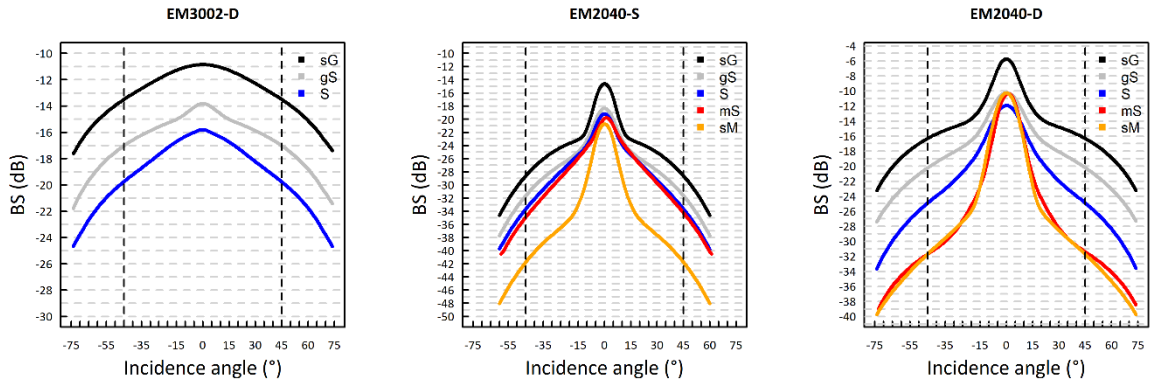


Figure 3.20 – Average backscatter strength (BS) angular profiles fitted using the GSAB model for each Folk category (Folk 5 scheme) and by echosounder model. Note that the inter comparison between echosounders is only relative as no absolute intercalibration was performed. The dashed black line indicates BS_{45}.

In Figure 3.21, the within class variability becomes apparent considering the overall envelopes. For the classes sand (S) and gravelly sand (gS) this reflects the exploratory analysis (boxplots) where the largest variability in backscatter values was observed for these classes, justifying the overlap with similarly sandy classes is expected. Accordingly, the sG class has the smallest variability envelope. It must be noted that besides the within class variability not captured by the classification scheme herein employed, part of the variability observed in the envelopes around the average curves for the three classes displayed in Fig. 3.21 will inherently be due to the multi-temporal nature of the dataset herein used. Depending on the hydrodynamic regime, tidally driven factors such as sediment transport in the water column and near-bed benthic boundary layer, short-term ephemeral deposition and erosion patterns and geometrical changes can influence the backscatter measurements up to ~4 dB at $BS_{45°}$ (Montereale-Gavazzi et al., 2019).

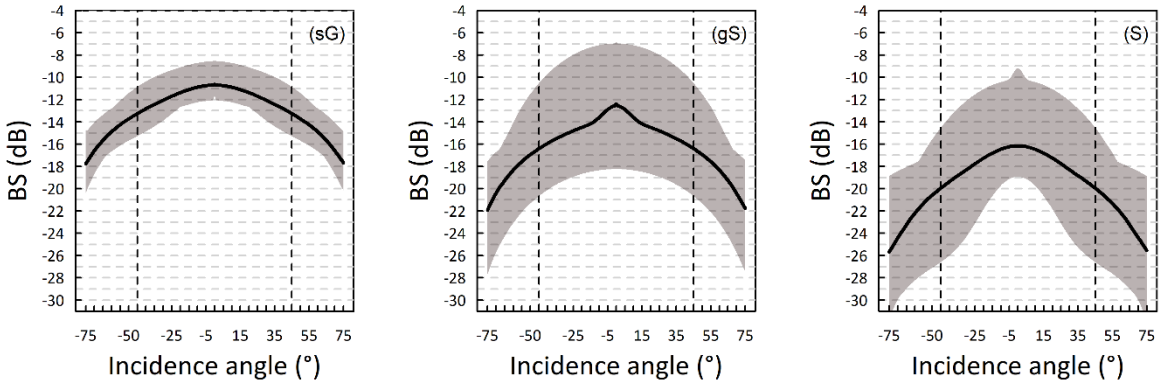


Figure 3.21 – Within-class envelopes of variability (minimum and maximum values of all angular profiles within a class) for the EM3002-D dataset. The dashed black line indicates BS_{45}.

3.11 DISCUSSION

The field of ASC into different themes (substrata/habitat) is developing at a fast pace through the application of MBES backscatter and bathymetry (Anderson et al., 2008). The ongoing trend towards the production of detailed (order of meters) and spatially-explicit resource/inventory mapping improves our understanding of underwater ecosystems, in space and time. The increasing amount of data becoming available and the need to inform policy-making with accurate information stresses the need to investigate semi-automated and repeatable objective classification routines that can form the basis of future comparative studies (i.e. monitoring of an area in respect to anthropic/natural changes; e.g. Rattray et al., 2013). In light of the available data, this investigation has mainly focused on three important aspects in the discipline of acoustic seafloor classification: i.e.: it has addressed (1) the relations between sediment and MBES backscatter, (2) the predictive accuracy of an unsupervised and a supervised classification routine available to seafloor mappers, and (3) the trade-offs between classification accuracy, classification scheme and discriminative ability of remotely-sensed data. Ultimately, the most accurate modelling approach (i.e. RF_{++} models) allowed a spatially-explicit mapping of the distribution of ecologically noteworthy gravel areas, providing fundamental information to guide future studies on poorly explored hard substrate benthic communities (e.g. for assessing their environmental status (e.g., De Mesel et al., 2017; in relation to marine aggregate extraction (e.g., Van Lancker et al., 2017), or wind-energy related infrastructural modifications (e.g., Degraer et al., 2016; Dannheim et al., 2019). Detailed identification of hard substrate areas advances further ecological applications as identifying areas suitable for benthic secondary producers (important for this part of the North Sea) poses the advantage of studying structural (benthopelagic coupling, recruitment, lifestyle) and trophic (food-web) links between pelagic and benthic systems and generally identifying candidate areas for conservation, possibly advancing the tutelage of the environment.

3.11.1 Sediment grain size and backscatter

Regarding the relationships between seafloor backscatter and sediment type, results are consistent with previous and recent studies and respect the general trend of higher reflectivity associated with increasing coarser sediment fractions and lower reflectivity with dominant finer ones (e.g. De Falco et al., 2010). The physical explanatory support of these relationships relates to the fact that backscatter is primarily controlled by three quantities (excluding angle of incidence and frequency): the impedance contrast between sediment and seawater, the roughness at the water-sediment interface and the volume inhomogeneity (inclusions in the sediment). As such, coarser sediments correspond to higher values of density and velocity (acoustic impedance) and to increased roughness, strongly linked to the grain-size content. The sediment interface roughness is itself related to the grain-size content, irrespective of its scale. Relatively flat sediment interfaces (i.e. lacking in topography of any sort), will have an intrinsic small-scale roughness given by the distribution of the grains. The scale of the roughness will thus increase with increasing coarseness of the sediment. Small-scale topography (i.e. centimetric and decimetric), such as micro-oscillatory ripples caused by tide and currents, have profound effects on the backscatter of high-frequency echosounders (e.g. Lurton et al., 2018; Montereale-Gavazzi et al., 2019). Siliciclastic sedimentary seafloors for instance, are prone to support current-induced relief whereas smoother and fluid-like sediments are not. The effects of volume (in-sediment) backscattering remains the least clearly explainable considering the available ground-truth data and generally, the fact that soft sediments can be highly variable in terms of the content of in-volume scatterers such as presence of gas bubbles, mineral inclusions and infauna (Rowden et al., 1998; Gorska et al., 2018). The variability observed in the sand (S) class may be indicative of these factors being left unaccounted by the sample description (note for example sample HB_VV5 in Fig. 3.9). As such, given the relatively weaker acoustic impedance contrast, soft sediments are prone to increased sound penetration, meaning that the backscatter level of one given grain-size distribution of fine sediment particles can be highly variable due to the occurrence of non-anticipated phenomena. Therefore, for the volume component, ambiguities are possible where unexpected results (inverse) could occur. For example, the presence of relatively fine (centimetres to decimetre) sand patinas covering gravel lag deposits could result in unexpected high backscattering levels; a factor observed in other studies (e.g. Calvert et al., 2015; Todd, 2005). Nevertheless, it is reiterated that for sandy substrates, it is expected that the penetration into the sediment for a 300 kHz operating frequency (as the data herein used) is limited to ~ 3 cm (Huff et al., 2008), comforting the assumption that the backscatter response would reflect the nature of the immediate surface. Ambiguities in the linearity and magnitude of increase of the relationship between backscatter and proportion of coarse sediment fractions may (in concert to biological, micro-scale roughness and geotechnical variables, all beyond the imaging capability of the sonar footprint and remaining unquantified from the ground-truth data) restrict the discriminative ability of backscatter in predicting

the full spectrum of substrate types (sensu Folk). In agreement with Gaida et al. (2018) and Snellen et al. (2018) it was shown that while backscatter generally linearly increases with increasing coarse sediment fractions and linearly decreases with increasing finer ones, considerable overlap exists between some of the Folk classes identified in the ground-truth data and a moderate plateau effect was observed with increasing percentages of coarse sediment fractions (particularly as the grain size approached the size of the wavelength i.e. 2-10 mm). This was particularly evident as the number of classes increased (considering the 8-class Folk scheme) and for the classes with a coarser (i.e. gravel/shell) sediment fraction. Therefore, for the coarser sediments a degree of intra-class dispersion (i.e. overlap) is expectable as a given grain size distribution will strictly control the impedance contrast, masking the contribution of the finer sediment fraction, irrespective of its nature. To overcome this limitation, though at the price of generalising the mapping product, class amalgamation was recurred to. Class amalgamation is a common limitation in acoustic seafloor classification using backscatter. Diesing et al. (2014) encountered similar issues using 300 kHz backscatter as mixed sediments could not be separated from coarser ones. Similarly, Gaida et al. (2018) and Snellen et al. (2018) equally could not find a 1:1 relationship between Folk and acoustic class: as a result, they also recurred to assigning multiple sediment classes to single acoustic classes. Fogarin et al. (2019), equally recurred to class amalgamation for the same reasons and using the same classification scheme. The D50 was shown to be a valuable predictor of sediment type on backscatter at the frequency herein used (i.e. 300 kHz); although this was limited to fine (in this investigation up to ~500 µm), well-sorted and mostly unimodal sediments (qualitatively referable to as plain/clean sediments). Following from Gaida et al. (2018), Snellen et al. (2018) and Buscombe et al. (2017), the plateau effect observed in Fig. 3.9C, and to a degree in Fig. 3.9B (i.e. sample HG_{14} with > 80% gravel content and > 50% 10 mm size fraction content) may be relatable to a transition and/or a mix of scattering regimes (i.e. from Rayleigh regime to discrete/geometric regime) when the mean sediment diameter/wavelength ratio exceeds 1 (around D50 = 5 mm for the frequency herein used). In this investigation (Fig. 9C), this is the case for the sediment fractions for which no magnitude of increasing correlation was observed, reaching a d/λ ratio of > 2 where the dominant percentage of sediment grains are ≥ 10 mm. A laboratory (tank) experiment by Ivakin and Sessarego (2007) found that a transition to negative correlation with backscatter intensity is expected when this occurs, pointing at the clear relationship between operating frequency and sediment type. It must be stressed however, that these laboratory experiments wherein this observation was made, were carried out under a well-controlled tank facility environment, using degassed, well-sorted, unimodal and artificially flattened granular sediment surface. As such, this observation would be valid under the assumption of seafloor homogeneity (not the case in this analysis: note the sample pictures in Fig. 3.9), where the grains of the sediment fractions are similar and evenly distributed around the median diameter. Organic content, micro-topography and bioturbation parameters which remain unaccounted by the sample description may also have an

influence on the acoustic returns. Besides the drawback of few observations to better explore these interesting relationships, investigating the sediment-acoustic relationship is of great relevance in view of unravelling the complexity of empirical/field measurements and improve our understanding of the discrimination performance, dictating the number of classes we are able to map and identifying the physical support explaining potential ambiguities of the acoustic response and subsequent limitations of acoustic seafloor classification based on a single frequency. In this regard, controlled laboratory and field hydroacoustic experiments (e.g. Ivakin, 2008, Williams et al., 2009), replicating complex seafloor scenarios encounterable in the operational environment, are needed to better elucidate these important factors.

3.11.2 Utility of harmonised multibeam multisource MBES data

Following Hughes-Clarke et al. (2008), Misiuk et al. (2018) and the recommendations set out in Lurton and Lamarche (2015) and Roche et al. (2018), the methodology employed here to produce a harmonised dataset of uncalibrated (i.e. relatively calibrated) backscatter was very successful at producing a seamless (i.e. continuous) character of the reflectivity across study areas and acquisition times. In this regard, relying on a stable backscatter reference area (such as the Kwinte Bank reference area; see Roche et al., 2018) poses the advantage of controlling the stability of the data recorded and allows obtaining empirically-derived BS offsets (by survey overlap). This is necessary to observe the continuity of substrate types and omit variations between-surveys and platforms (for example the source levels of different echosounders). This process allowed exploitation of the valuable backscatter dataset at once, as oppositely to studies where the analyses were carried out classifying each survey separately and merged via manual procedures (Lacharité et al., 2018). In predicting the distribution of substrate over spatial extent similar to that in this investigation (~ 135 km^2), Misiuk et al. (2018) found that directly exploiting the harmonised dataset increased accuracies of model predictions. As such, harmonisation of MBES datasets improves model performance as well as reducing operator/machine times. The value of the cross-calibration propagation could be enhanced by cross-calibrating the echosounders of the platforms with an echosounder with an absolute calibration (thus using the absolute data as the nominal reference). Recently, methodologies based on use of natural reference areas and absolute calibrated data have been proposed (Eleftherakis et al., 2018; Roche et al., 2018). With increasing volumes of data being systematically acquired, merging of multisource surveys to produce large geographical coverage maps of substrate types will increasingly be needed. In this regard relying on natural reference areas for cross-calibration propagation (Eleftherakis et al., 2018; Roche et al., 2018; Weber et al., 2018) proves to be a successful way forward and a pragmatic emerging approach for hull-mounted systems. In this investigation, a simple inter-calibration based on offsetting of compensated backscatter imagery to a reference

value at 45° was implemented and the full inter-calibration of the systems (i.e. across all angles based on a reference calibrated model) remains work in progress.

3.11.3 Impact of classifier on model performance

Knowing which classifier performs best on the available data can be a challenging task (Foody et al., 2007). With this, the need comes to test multiple approaches so that the best performing one can be selected for future applications (only few published comparative studies have been proposed: Ierodiaconou et al., 2011; Stephens and Diesing, 2014; Calvert et al., 2015; Montereale Gavazzi et al., 2016). The primary goal of this investigation was to evaluate the predictive performance of an unsupervised and a supervised approach in capturing the distribution of substrate type via the integration of multi-source MBES and ground-truth data. The main reason for such an investigation is the need to identify an objective and repeatable routine that can be used in future mapping and monitoring applications and that can maximise the predictive accuracy of the thematic model produced. Unsupervised k-means and supervised Random Forest classifiers were chosen to test the performance on three sets of input data: backscatter alone, backscatter and bathymetry and backscatter, bathymetry and a further set of selected morphometric and textural derivatives. The accuracy of the thematic models produced was statistically assessed using the contingency table (or confusion matrix), allowing the cross-tabulation of predicted and observed instances and the subsequent derivation of global and by-class accuracy measures (Foody, 2002; Liu et al., 2007).

3.11.4 Unsupervised approach

Unsupervised k-means was implemented to test the utility of using backscatter on its own, being the most relevant proxy of sediment type (as reported in a range of studies; e.g. Collier and Brown, 2005; De Falco et al., 2010; Ferrini and Flood, 2006; Goff et al., 2000, 2004), without recurring to rather complex classification routines requiring increased amounts of data types (see Stephens and Diesing, 2014). Unsupervised clustering was selected due to previous research reporting promising classification accuracies based on simple clustering of backscatter alone (e.g. Tęgowski, 2005; Fezzani and Berger, 2018). For example, Montereale-Gavazzi et al. (2016) successfully applied Jenk's optimization clustering (Jenks, 1967) (an algorithm in the same clustering family as k-means) to predict the distribution of five benthic substrate classes in a highly heterogeneous tidal channel of the Venice Lagoon (Italy). In this study, where several classifiers were compared, it was shown that a simple clustering approach produced the most accurate results when compared to other machine learning classifiers (> 80 and 70 % in overall accuracy and kappa statistic). This is in accordance with recent mapping studies in the northern Adriatic Sea (Fogarin et al., 2019) and in Australian waters (Hasan et al., 2016). Furthermore, Montereale-Gavazzi et al. (2017) found that a relatively simple k-means clustering classification approach compared well with the results obtained

from a Random Forest modelling, producing small differences (i.e. < 10 % difference between classifiers) of three substrate classes within the same study area of this investigation (a restricted part of the Hinder Banks). However, in the current investigation and for this study area, this classification performed relatively poorly (particularly with increasing number of classes). Lower scores of global and by-class accuracies were found, insufficient to provide adequate information for any further application.

An important misclassification occurred particularly between predominantly muddy areas and top of sandbanks due to the similarity of the backscatter mosaic values, the former being classified as sM (see Model performance in the Results section). A further limitation posed by using unsupervised routines is that of defining objectively (i.e. statistically identifying) the number of clusters. In this investigation the WGSSD and the silhouette coefficient metrics were tested as in for example Eleftherakis (2013) and Snellen et al. (2018). The WGSSD gave an output of three to five clusters as the optimal solution whereas the silhouette suggested three clusters as the optimum although sufficient class separation (i.e. silhouette coefficient > 0.55) was also found for a five-cluster solution. It must be noted that finding the number of clusters in backscatter data is not a trivial task given its random/noisy nature leading to a natural overlap of the ranges in values defining the clusters, making the statistical identification of the correct number of classes a very challenging aspect (as noted in Snellen et al., 2018 and Fezzani and Berger, 2018). Nonetheless, in the present study it was shown, via the interpretation of boxplots, that three and five-class solutions were reasonably discernible in the backscatter dataset (good separation of the median values and maintaining a linear trend from finer to coarser sediments). Still, improvements of the unsupervised approach may be possible. The objectivity of the k-means classification could be improved by first conducting a Principal Component Analysis (PCA) on a set of backscatter derivatives (e.g. mode, minimum, maximum, median) to, in turn, cluster the combination of Principal Components (PCs) which explain most of the variance. This approach has been used by, amongst others, Eleftherakis (2013) where it was however noted that given the very high correlation between the mentioned backscatter features (as equally observed in the Feature selection of this investigation), a classification based on each feature independently would be very similar, therefore not effectively enhancing the discriminative ability of backscatter alone. This implies that in this investigation, the same “confusion” between top of sandbanks and predominantly muddy areas (as discussed in the Results section), would not have been overcome by this approach. Due to this reason, it was decided to undertake a classification integrating bathymetric and textural derivatives, using the RF classifier which has shown remarkably promising results in recent seafloor mapping research (Hasan et al., 2012; Stephens and Diesing, 2014; Ierodiaconou et al., 2018).

3.11.5 Supervised approach

The routine set up using the RF approach had several advantages compared to the simple unsupervised routine and produced very accurate predictions ($RF_{++EUNIS}$; A = 0.85, K = 0.76, NIR = 0.47 and RF_{++Folk}; A = 0.74, K = 0.63, NIR = 0.37) which are of great relevance for an array of further spatial and ecological applications. Firstly, the algorithm benefits from an internal form of feature importance selection (the Boruta RF wrapper function), allowing the production of simpler models (Stephens and Diesing, 2014). Unsurprisingly, backscatter and bathymetry, followed by slope, roughness and Moran autocorrelation, were identified as the most relevant variables whereas textural indices computed from the backscatter and the topographic position index were found to be irrelevant. A further advantage of the supervised RF routine is that the algorithm makes no assumption about the distribution of the underlying data and it therefore outperforms algorithms based on assumptions of normally distributed input layers such as the Maximum Likelihood Classifiers (Hasan et al., 2012; Ierodiaconou et al., 2011) and the herein tested k-means clustering which has the inherent tendency to find clusters of similar magnitudes and spherical shape (Patil and Vaidya, 2012; Lu et al., 2017). Moreover, RF is relatively insensitive to overfitting as the predictions are drawn from aggregation of multiple randomly constructed decision trees (via the bootstrap sampling procedure), which reduces the variance of the predictions. The possibility to tune the model prior to its implementation has also the advantage to render the routine more objective and the output more accurate than relying exclusively on default parameters. For example, tuning of the parameters ntree (number of trees grown in the overall forest) and mtry (number of randomly selected predictors used in the tree construction phase) enables a considerable improvement of the predictive performance. In this regard it has been demonstrated that the smaller the number of trees grown (particularity below 300 trees) the poorer the performances given that the chance of each variable being present in several trees reduces (Gazis et al., 2018). Indeed, it is very promising to find that substrate and benthic habitat maps produced by the same RF algorithm in other geographical areas and based on similar input of predictor layers, leads consistently to higher performances compared to former classifiers used in the seafloor mapping literature (Hasan et al., 2014; Hasan et al., 2012; Ierodiaconou et al., 2018; Porskamp et al., 2018; Stephens and Diesing, 2014).

3.11.6 Impact of prescribed classification scheme on model performance

At the core of acoustic seafloor classification lies the selection of a substrate/habitat classification scheme (Strong et al., 2018). The prescribed scheme dictates the ultimate way a seafloor area is thematically represented and hence the type of information that will be communicated from mappers to environmental managers, policy makers and stakeholders. Choosing the adequate scheme is not trivial and can have profound effects on model accuracy and the type of information it contains. In this study, two substrate classification schemes commonly used at the European

level were investigated (Diesing et al., 2014; Kaskela et al., 2019). Using the RF supervised classifier very good results were obtained for both classification schemes. In this regard, it was noted that as the number of classes increased, the model accuracy decreased. As noted in similar studies by Montereale Gavazzi et al. (2016) and Porskamp et al. (2018), this effect is likely due to the increased complexity of the classification task, compounded by the decreasing number of training samples per class and leading to the establishment of “rarer” classes. This was particularly evident when the RF algorithm was tested on the finest level of Folk classes identified (n = 8 categories), which produced the lowest accuracies. Larger amounts of training samples per category are inherently associated to increased performance of this algorithm (Millard and Richardson, 2015). Besides, the sampling effort on the acoustic surveys used in this investigation was primarily skewed towards the poorly explored hard substrate areas. Considering the high costs of ground-truthing, sample datasets could benefit from existing data (i.e. legacy datasets), though this would pose the limitation of training and validating classifications with samples that may not necessarily be representative of the time of survey. Clear trade-offs exist between accuracy of the model, the number of classes mappable in respect to the available ground-truth data and the discriminative ability of the remotely-sensed data; mappers will have to make a choice between the downside of having a fine-scale detailed map with low accuracy and a highly accurate map with fewer categories. Overall, the classification schemes herein used enabled a satisfactorily supervised prediction of broad and finer-scale patterns of sediment distribution for the areas surveyed. Regarding the discrimination potential of backscatter, new insights could come in from the use of recent developments in multi-frequency backscatter measuring sonar systems. Novel research in this application (e.g. Buscombe and Grams, 2018; Feldens et al., 2018; Gaida et al., 2018) provides an indication of the added value of accounting on the “third dimension” of backscatter (i.e. the subsurface volume backscattering – acoustic penetration), currently treated as a two-dimensional process.

3.11.7 Modelling of angular responses

Despite the regrettable lack of a proper inter- and absolute calibration of the angular responses (i.e. applying dB offsets obtained from a calibrated reference model across all angles - a currently ongoing long-term project requiring the involvement of several parties – Lurton and Lamarche, 2015), their modelled profiles proved as an optimal validation technique of the classes identified, providing the means to confirm their physical and sedimentological differences. As observed in Fezzani and Berger (2018), envelopes of variability for each category are visually widespread, reaching quantities > 10 dB at 45° and suggesting that a finer classification scheme would be feasible. However, this would require a significant increase of the ground-truth effort to accommodate the observed within-cluster variability. This type of empirical modelling is particularly applicable given that, to be accurate, theoretical models of angular response may require too many input parameters for being practical,

although never describing satisfactorily the very complex nature of the acoustic interaction with heterogeneous marine sediments. Furthermore, the complexity of marine sediments (i.e. layering, heterogeneity of the sediment bulk properties and/or roughness) which remains unaccounted in the description of the ground-truth data, makes the a priori application of a unique geophysical model irrelevant. This is exacerbated at a regional scale, where a model built for a specific seafloor area (such as the APL model in Jackson et al. 1986 - see the Introduction section) will be inapplicable. The application of a more descriptive and generic model (such as the GSAB) is highly valuable as it provides a simplified, yet physically meaningful description of the sediment type and can thus be applied to discern between the main substrate types (such as those identified in the current investigation).

3.11.8 Limitations: sources of error and possible improvements

A set of general limitations to the mapping approach and the underlying data should be considered in future applications, particularly in respect to the ground-truth sampling. In this investigation, samples were acquired within 48 h from the acoustic survey, hence they are closely representative of the spatio-temporal status of the seafloor at the time of survey: this is a considerable improvement in respect to the production of maps that rely on legacy datasets (e.g. Stephens and Diesing, 2014), prone to the propagation of spatio-temporal errors in the classification process (depending on the temporal variability caused by hydro-meteorological conditions). The time lag between the acoustic survey and the ground-truth sample acquisition can be a more significant source of error within highly dynamic nearshore areas while it can have a lesser impact in more stable and offshore environments (Montereale-Gavazzi et al., 2019). Nonetheless, it must be noted that there are fair disagreements between the sampling size (this varied considerably across the sampling gears herein used; the largest being the Hamon Grab with ~ 0.1 m^2) and the grid resolution of the acoustic data (25 m^2). It is reasonable to assume that significant small-scale variability will be present within a seafloor portion of 25 m^2 (Fig. 3.22, 3.23). The scale of predictor variables is a topic that equally deserves future attention as it may have repercussions on model accuracy: the relative importance of predictors in respect to given substrate classes has been found to vary considerably with varying grid resolution (e.g. Misuk et al., 2018); while this was not investigated here, it is reasonable to expect that computation of finer scale derivatives (i.e. using 1 m pixel size, feasible in the coastal areas) may improve the prediction of fine-scale, heterogeneous patterns such as distinguishing between the patchily distributed sM and mS classes. Besides, these classes had the smallest number of samples, also possibly justifying the comparatively poorer performance of their predictions. Errors may also arise due to a mismatch in the positional accuracy of the samples and the acoustic data, resulting in potentially misleading information (especially for highly heterogeneous seafloor areas). Use of a buffer (as in this study) is recommended, though its adequacy will depend on the patterns and scales of seafloor heterogeneity. Improvements in positional accuracy of the samples could

come in from the integration of acoustic tracking devices on the sampling gear (Coggan et al., 2007) whereas improvements to ensure the spatial homogeneity of given seafloor areas could come in from the acquisition of replicate samples (although increasing the sampling effort in terms of cost, timing and labour intensity). Furthermore, it is important to consider the representativeness of the retrieved sample; both in terms of the sampling gear and the analytical procedure employed. For example, the Hamon grab sampler retrieves a highly disturbed sediment sample without preserving the integrity of the strata being sampled. This has implications when describing the coarse sediment fraction which may not be representative of the water-sediment interface (i.e. the classification target). Besides, sampling of hard substrate areas must rely on adequate gears for an effective sampling. A further, potential source of error/misinterpretation, is the subsample used in the laser granulometric analysis (in this investigation using the Malvern Mastersizer instrument): this may not be representative of the surface sampled by the gear, as well as of the broader extent from which the estimates of backscatter are derived (values from mosaics or angular responses). In this context, the sand:mud ratios that dictate the Folk type class could be unrepresentative (again depending on the spatial homogeneity of a given area). Improvements could come in from the analysis of subsample replicates, consequently averaged and possibly leading to more robust estimates of the sediment type. A further fundamental issue is the observer-bias in the sample description phase. In this study, the estimates of the fine- to medium sand fraction (0.063 – 1 mm) can be considered as accurate and objective, being derived from a Malvern Mastersizer 3000 analysis (the accuracy of this instrument is in the range of 1 % - Malvern Panalytical, 2019), whereas the subjectivity of the estimates of the coarser sediment fraction (> 1 mm) will depend on the observer judgment, hence on expertise and on the quality of the available photographic material collected on board during the sampling phase of the investigation. Here, improvements could come in from the standardisation of photographic protocols, for example acquiring planar pictures of the sample surface, consequently allowing automated image analysis (see Chapter 1 – Ground-truth acquisition and processing). In this regard, in a study comparing the inter-observer agreement at describing four to six seafloor categories from video and photographic ground-truth data (thus approximating to the number of categories in the present investigation), Rattray et al. (2014) found strong overall agreement accuracies consistently > 75 % between three observers of varying training and expertise backgrounds.

While videographic sampling protocols remain underdeveloped in Belgian mapping efforts, they are very promising as the field of view approaches the gridded acoustic resolution and the potential for automated image analysis, deriving for example the percentage of coarse sediment, is greatly enhanced. In this investigation, videographic sampling experiments have been conducted during the ST1817 RV *Belgica* campaign (Fig. 3.1B, Table 3.1 – see campaign report available at: https://odnature.naturalsciences.be/downloads/*Belgica*/campaigns/reports/re2018_17.pdf), partly dedicated to the exploration of the northwesternmost offshore area of the Belgian jurisdictional maritime area. The experimented acquisition and analysis

protocols remain suboptimal although they showed promising results (Fig. 3.22, 3.23 and Ch. 2 Fig. 2.17). Furthermore, it must be noted that this technique is entirely non-invasive, fully avoiding impacts on the benthic biota and on the physical structure of the seafloor. This is contrary to certain destructive techniques (e.g. trawled Gilson dredge sampler – Gilson 1907; Houziaux et al., 2011) currently in use and which should be dismissed, particularly within Habitat Directive and conservation areas (Assis et al., 2008), and in the context of MSFD monitoring of the highly pressured (i.e. from fisheries) offshore gravel beds of Belgium. Furthermore, it must be noted that while epibenthic species can be efficiently sampled using dredge sampling gears, the position of the retrieved biota will be limited to the extent of the trawl, hindering the determination of exact locations in the surveyed areas which has implications for subsequent modelling applications. Figure 3.22, reports details of video drop-frame acquisitions in the Northern exploration area, indicating the potential of this kind of data. Noticeably, the ground-truth frames match well with the lower (predominantly sandy – 70-90 % range) and higher (predominantly gravel – 50-70 % range) reflectivity patterns observable in the backscatter image. Figure 3.23 displays a set of still frames acquired within this area, clearly pointing at the potential for identification of biota (at least considering morphospecies and phylum level). In quantifying the seafloor interface parameters which are of interest in backscatter studies, the appreciation of the volume components of the sediment matrix remains limited by this approach, although the quantification of small-scale roughness/morphology (Fig. 3.23G), influencing the backscatter response becomes possible (Ferrini and Flood, 2006; Lurton et al., 2018).

Empirical investigations applying multivariate statistical analyses to MBES backscatter have shown great potential in enhancing the understanding of the combined drivers of the acoustic response (e.g. Ferrini and Flood, 2006). It remains true that the past few decades have seen a remarkable technological breakthrough of the hydroacoustic remote sensing technology, whereas the seafloor sampling gears, critical in view of acoustics interpretation, have remained largely unchanged since historical times. Sediment Profile Imagery (e.g. Solan and Kennedy, 2002) has been a noteworthy progress in this regard, allowing the undisturbed observation of the sediment volume inhomogeneities. Combined use of gears operated from vessels with a dynamic positioning system (DPS – having positional and heading accuracies of 1-3 m and 1° respectively such as on RV *Simon Stevin*) may be a way forward promoting a more holistic understanding of the sediment being sampled and of the resulting seafloor backscatter response. An example of this is given in Figure 3.24 where SPI and Reineck box core samples where coincidently acquired at five sample locations (corresponding angular response backscatter curves are shown). Clear insights into the origins of the acoustic response can be achieved: herein related to the immediate water-sediment interface, related to the presence of bioclastic detritus of varying magnitudes, drastically modifying the water-sediment impedance contrast and roughness.

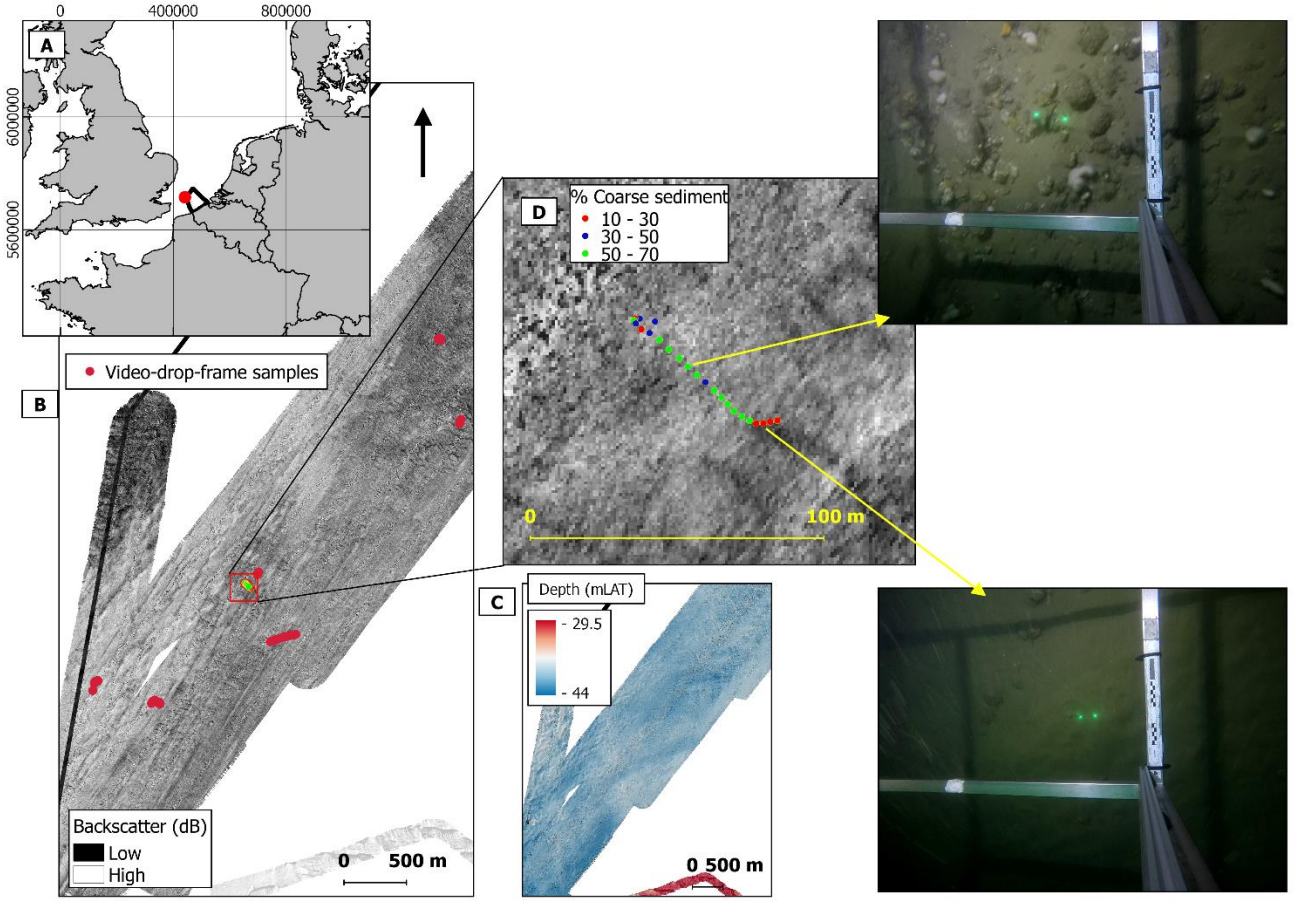


Figure 3.22 – Cartographic summary of the MBES and videographic data acquired in the Northern exploration area during the ST1817 RV *Belgica* campaign. A) Location of the Northern exploration area within the Belgian Part of the North Sea. B) Backscatter mosaic with videographic sample locations over imposed. C) Bathymetric map of this survey area. D) Detail of a video-frame transect classified into percent cover of coarse sediment fraction. Examples of the imagery are displayed on the right-hand side. The laser pointers in the images are 9 cm apart. The gear used includes a GoPro, a Hugyfot Arius 1500 lm Video lighting torch and an additional QUDOS 400 lm torch. The percentage covers were extracted from the still frames via an automated routine combining pattern recognition and supervised classification utilities of the WEKA Trainable Segmentation Tool (Arganda-Carreras et al., 2017); a dedicated image classification plugin of the ImageJ image processing software (Rasband, 2012).

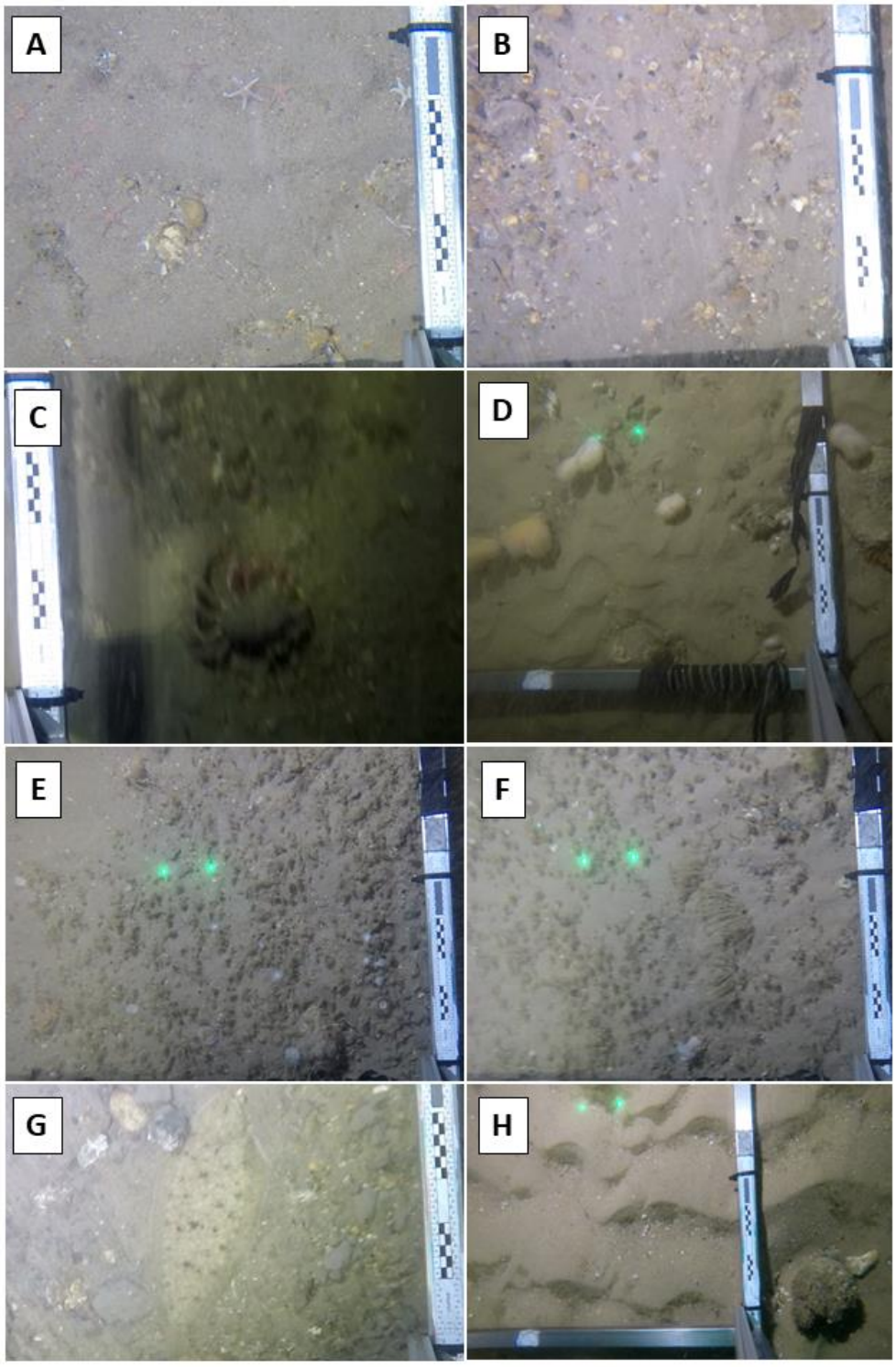


Figure 3.23 – Examples of still frames obtained by the video-drop-frame sampling approach in the Northern Exploration and Hinder bank area. A-B) visible Asteroidea, C) Crustacea (*Maya* sp.) roaming on a gravelly substrate. D) Soft corals (*Alcyonium* sp.) bio encrusting cobbles and boulders. E-F) Highly biogenic substrate including Anthozoans, crustaceans and *Nemertesia* sp. G) Hard substrate promoting occupancy by flat-fish. H) Noticeable micro ripples (λ = ~15 cm) and a boulder encrusted by bryozoans.

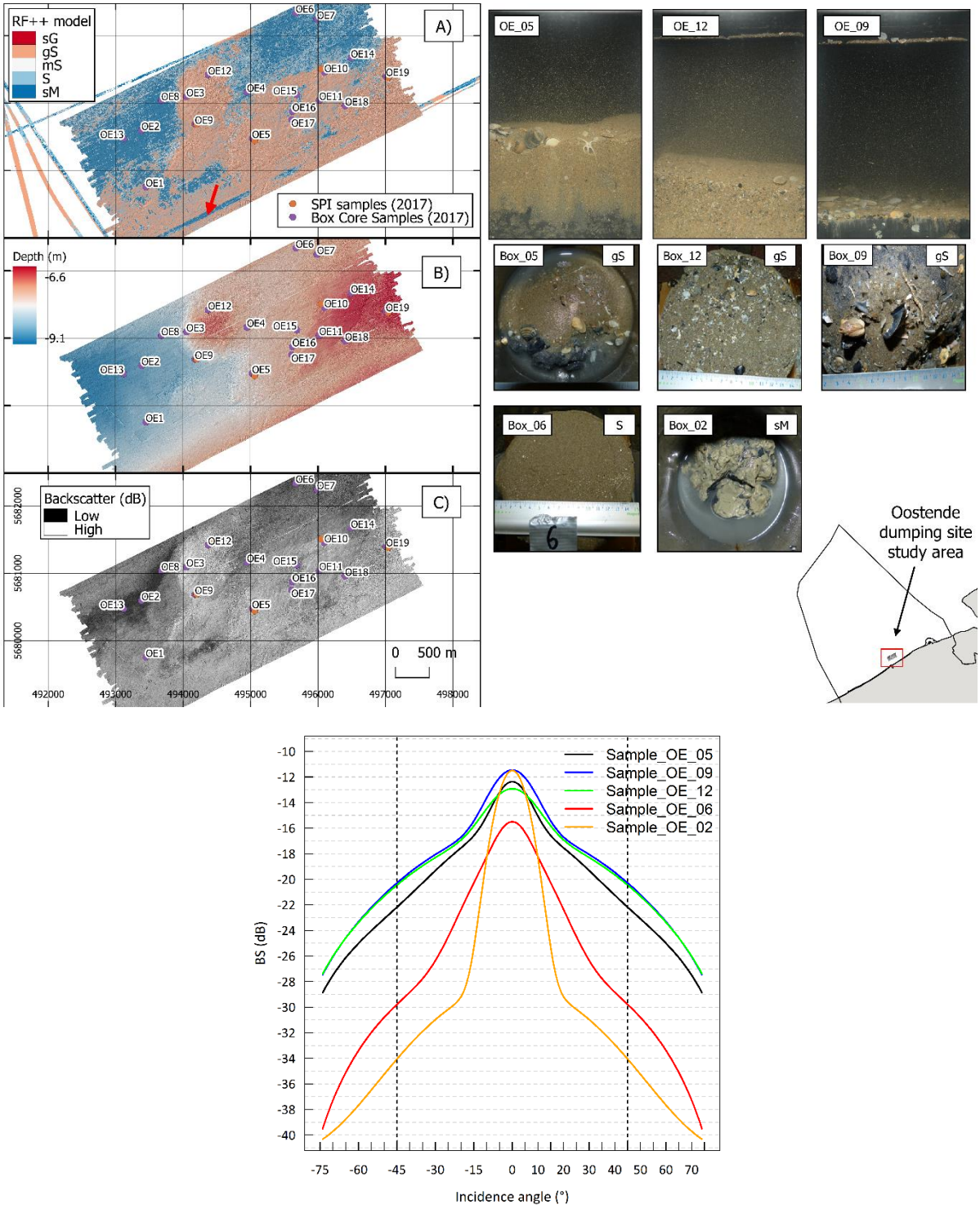


Figure 3.24 - Hydroacoustic and ground-truth data for the Ostend disposal ground of dredged material study area. The location of the study site within the BPNS is displayed in the bottom right corner. A) Classification of this area by the RF$_{++Folk}$ model, B) RTK-corrected bathymetry and C) co-located compensated backscatter imagery. Ground-truth locations are superimposed on each grid. Pictures of five selected box-core samples (three of which have a co-located SPI sample) are displayed on the right-hand side. Sample labels match between maps and sample images for interpretation. Below are the modelled angular response curves of the five samples presented. Note the remarkable similarity (i.e. ~1 dB at 45°) between OE_5, 9 and 12 despite the apparent differences in depth of the sand layer: it is clear that few percentages of shell/gravel, drastically modifies the acoustic response by dictating the impedance contrasts and roughness components.

3.12 Conclusions

Seafloor mapping through hydroacoustic and ground-truthing approaches are developing at a fast pace, stimulated by a diversity of marine management relying on spatially explicit information for the acknowledgment and monitoring of the submerged environment (e.g., Marine Spatial planning, Marine Strategy Framework Directive, Ecosystem-Based Management). MBES have become the mainstream tool in this realm and increasing volumes of such data are being collected by a multitude of initiatives. This stresses the need to test and identify efficient mapping approaches, transforming the hydro-acoustic information into products that are useful at scales relevant for management. In this regard, investigating approaches producing accurate baseline mapping efforts are indispensable towards the development of seafloor mapping strategies and future monitoring applications.

Increasing awareness on map production objectiveness, repeatability and communication of accuracy is needed to optimise the integration with existing sediment maps; moving away from subjective classification procedures which currently form the basis of most national and international seafloor mapping and “data-ingestion” initiatives (e.g. EMODnet - lacking accuracy metrics and limited to confidence scores – i.e.: non-statistical but rather driven by expert interpretation). This requires testing the effectiveness of semi-automated data-integration and classification routines, a topical issue which has received increasing attention over the past few years. Furthermore, ground-truthing of backscatter needs developing and access to a wider array of sediment (and biologic) parameters, enhancing the understanding of seafloor-acoustic relationships in the operational context as well as enhancing the information derivable from remote sensing.

This research work focused on the predictive mapping of seafloor surficial sediment distribution via the integration of MBES and ground-truth data acquired in the Belgian Part of the North Sea. In the light of the acquired datasets, sediment-acoustic relationships were investigated, and two classifiers and two substrate classification schemes were tested for their predictive accuracy, providing useful insights into the trade-offs of the thematic representation of the seafloor. Both classifiers were found to decrease in predictive accuracy with increasing number of classes due to the combined effect of an increased complexity of the classification task and the availability of samples per category, leading to the establishment of increasingly rarer classes. Additionally, modelling of angular response backscatter angular responses was carried out, maximising the information obtainable from MBES backscatter measurements.

Datasets acquired during eight oceanographic campaigns (covering approximately 150 km^2 and ground-truthed by various approaches) were successfully merged by applying a simple cross-calibration propagation based on mosaicked backscatter and survey overlap on a stable natural reference area, producing a seamless map of the surveyed areas and predicting the distribution of benthic substrates reaching accuracies > 80 %. The emerging concept of natural

reference areas is a keystone and pragmatic solution requiring development at the European level to improve merging of various mapping efforts.

Sediment-acoustic relationships were observed relating backscatter ranges to variation in grain-size parameters (median diameter and percent weight of finer and coarser sediment fractions). Fundamental empirical evidence was obtained, evidencing the implications these have for assigning frequently used sediment classification schemes, such as Folk-type classes, to single- and high-frequency backscatter data. The shortcomings of the backscatter discriminative ability within heterogeneous coarse sediments must be accounted for during the data integration phase of mapping studies. Class aggregation is required to force the prescription of such a classification to the hydroacoustic data. Investigations based on multi-frequency systems, similarly to the terrestrial remote sensing realm, may be a way forward enhancing the discriminative ability of seafloor backscatter and provide new insights.

Under these constraints, the classifications showed the ability to discern between a sufficient number of classes that could be seamlessly predicted over the entire acoustic dataset. This implies that where repeat surveys and classification routines are put in place (and kept to a rigorous standard), areal and physical changes can be quantified. In line with current research, the supervised machine learning routine set up using the Random Forest classifier holds great potential for data integration and is found to supersede former generation classifiers.

Appendix A – Raw confusion matrices

Raw confusion matrices are reported in Table 3.A1. User and Producer by-class Accuracies are given in the last two righthand columns.

Table 3.A1 – Raw confusion matrices for the classifications performed. The first three matrices are for the five Folk classes classification scheme, RF ++ model, RF simple model and K-means classifications. The last three matrices are for the three EUNIS Habitat Level III classification schemes in the same order. Bold and italicised values are the diagonal entries (i.e. correctly classified). Other values are the from-to entries between predicted and observed instances. Column and row totals are reported below and to the right of each matrix respectively. User and Producer accuracies are reported in the last two columns. Producer Accuracy refers to: the number of correctly classified samples of class x divided by the total number of validation samples of class x (Column Totals). The obtained value relates to the probability that a ground-truth sample is correctly classified. This relates to the error of omission, meaning that instances which have not been correctly classified as x were omitted from the allocation to the correct class (i.e. how well a specific seafloor area can be classified?). User accuracy (also referred to as reliability refers to: the number of correctly classified samples of class x divided by the total number of samples that were allocated to category x (Row Totals). The obtained percentage indicates the probability that a sample classified on the map actually represents the category on the real ground. This relates to the error of commission, meaning the reliability of the thematic model (i.e. how well does the map produced represent reality? – Story and Congalton, 1986).

RF++ 5 Model Class	sG	gS	mS	S	sM	Row Totals	User	Producer
sG	**15**	1	0	1	0	17	0.88	0.88
gS	2	**16**	0	1	0	19	0.84	0.8
mS	0	0	**1**	1	1	3	0.33	0.5
S	0	3	1	**6**	2	12	0.5	0.6
sM	0	0	0	1	**2**	3	0.66	0.4
Colum Totals	17	20	2	10	5	54		

RF simple 5 Model Class	sG	gS	mS	S	sM	Row Totals	User	Producer
sG	**15**	0	0	0	2	17	0.88	0.88
gS	1	**15**	0	3	0	19	0.78	0.88
mS	0	0	**1**	0	2	3	0.33	0.33
S	1	2	1	**4**	4	12	0.33	0.57
sM	0	1	1	0	**1**	3	0.33	0.11
Colum Totals	17	18	3	7	9	54		

KM 5 Model Class	sG	gS	mS	S	sM	Row Totals	User	Producer
sG	**13**	2	0	2	0	17	0.76	0.81
gS	2	**8**	1	8	0	19	0.42	0.73
mS	0	0	**1**	0	2	3	0.33	0.43
S	1	1	5	**0**	5	12	0	0
sM	0	0	0	1	**2**	3	0.67	0.33
Colum Totals	16	11	7	11	9	54		

RF++ 3 Model Class	Mud	Sand	Coarse	Row Totals	User	Producer
Mud	**6**	1	0	7	0.85	0.66
Sand	3	**21**	2	26	0.8	0.91
Coarse	0	1	**14**	15	0.93	0.87
Colum Totals	9	23	16	48		

RF simple 3 Model Class	Mud	Sand	Coarse	Row Totals	User	Producer
Mud	**4**	3	0	7	0.57	0.66
Sand	2	**24**	0	26	0.92	0.75
Coarse	0	5	**10**	15	0.66	1
Colum Totals	6	32	10	48		

KM 3 Model Class	Mud	Sand	Coarse	Row Totals	User	Producer
Mud	**5**	2	0	7	0.71	0.35
Sand	9	**9**	8	26	0.34	0.69
Coarse	0	2	**13**	15	0.86	0.61
Colum Totals	14	13	21	48		

Appendix B – Partial dependence plots of the RF_{++Folk} model and selected variables

Random Forest is a powerful classification algorithm, yet it makes the identification of functional relations among predictors and prediction outcomes difficult. Partial dependence plots can be used to capture insights of these relationships. These are depicted in Figure 3.B1 for the five most important predictor variables used to grow the forest (ensemble of trees) of the RF_{++Folk} model. The relationship between predictor variable and the substrate class selection by the RF model is graphically examined. The partial dependence measures the marginal effect of a given variable on the class response while averaging out the effect of other variables in the classification process (Liaw and Wiener 2002). The datapoints in the partial dependence plots refers to the average percentage vote in favour of the substrate class over all observations and at given fixed levels of the explanatory/predictor variable under consideration. Therefore, the plot provides insights into the relative importance of these selected predictors to the selection of a substrate class over the entire range of the predictor.

Interpreting the plots, the five predictor variables have little influence on the selection of the sM class, with values of backscatter in the range ~ -45 to -38 dB (lowest bound of the backscatter range) having the largest influence on the prediction of this class. The mS class is strongly anti correlated to the sG class and its prediction is highly influenced by lower values of backscatter (range -45 to -32 dB), bathymetry < -20 m, lowest values of roughness and slope (i.e. lack of terrain complexity) and higher values of spatial autocorrelation (i.e. indicative of similar values clustering together on the grids). The S and gS class experience similar influences by all predictor variables, denoting their similarity. Expectedly, higher reflectivity (backscatter intensity) has a greater influence on the prediction of these classes. Bathymetric variation appears to have a stable influence on these classes as much as slope. Contrarywise and expectedly, higher values of roughness (i.e. more complexity) and lower values of spatial autocorrelation (i.e. denoting dissimilar values clustering together on the grids), positively influence the prediction of these classes. This pattern reflects the high heterogeneity and patchy distribution of these classes. The sG class is strongly positively influenced by the upper bound of backscatter intensity in the range -32 to -25 dB. The bathymetric range influences the prediction of this class in its deepest range, reflecting the predominance of this class in the deepest portions of the overall study area (depth range -20 to -40 m). Intermediate values of roughness ~ 0.5 to 1 and slope > 2° influence positively the prediction of this class whereas intuitively, low spatial autocorrelation contributes to this class prediction.

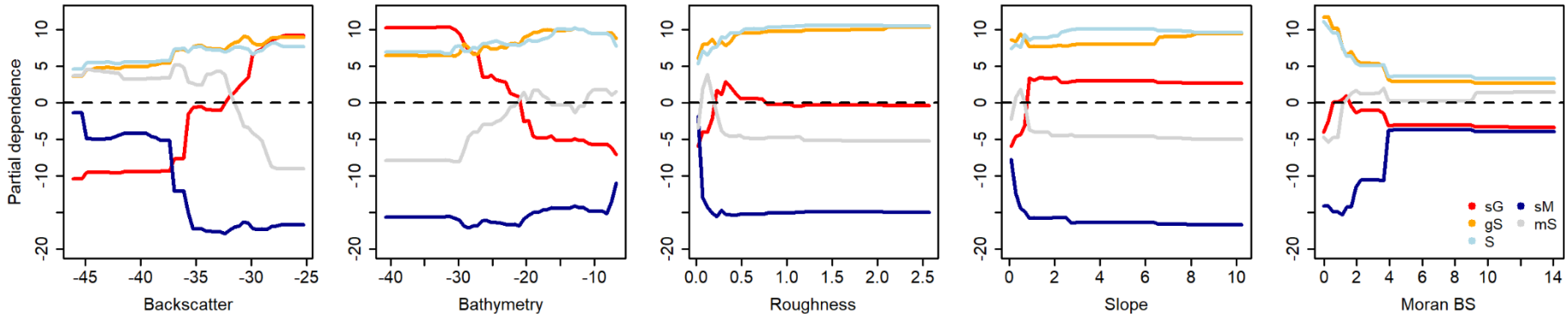


Figure 3.B1 – Partial dependence plots for the RF_{++Folk} model and the five most relevant variables identified by the feature selection routine, used to build this model. Left to right are plots of partial dependence on Backscatter, Bathymetry, Roughness, Slope and Moran Backscatter Autocorrelation. Vertical axes %. Predictor variables on the x axis. Trends are colour coded according to the predictive maps reported throughout this document (see legend bottom right).

Appendix C –RF$_{++Folk}$ 8 classes model

Herein, the results obtained by applying the RF++Folk modelling approach to the 8 classes Folk scheme are reported. Figure 3.C1 and Table 3.C1 report the maps acquired by this method and the raw confusion matrix, along with by-class accuracy metrics, respectively. Overall, this model scored an Accuracy of 58 % [0.44-0.71 95 % CI], a Cohen's kappa of 47% and a NIR = 0.36. While the NIR remains below the overall Accuracy and kappa metrics, the model is found to perform poorly compared to the 3 and 5 class solutions. This result aligns to the notion that increasing the number of classes increases the complexity of the classification task. Unsurprisingly, classes having the weakest by-class accuracies are those that were found to be considerably overlapping in the exploratory data analysis section.

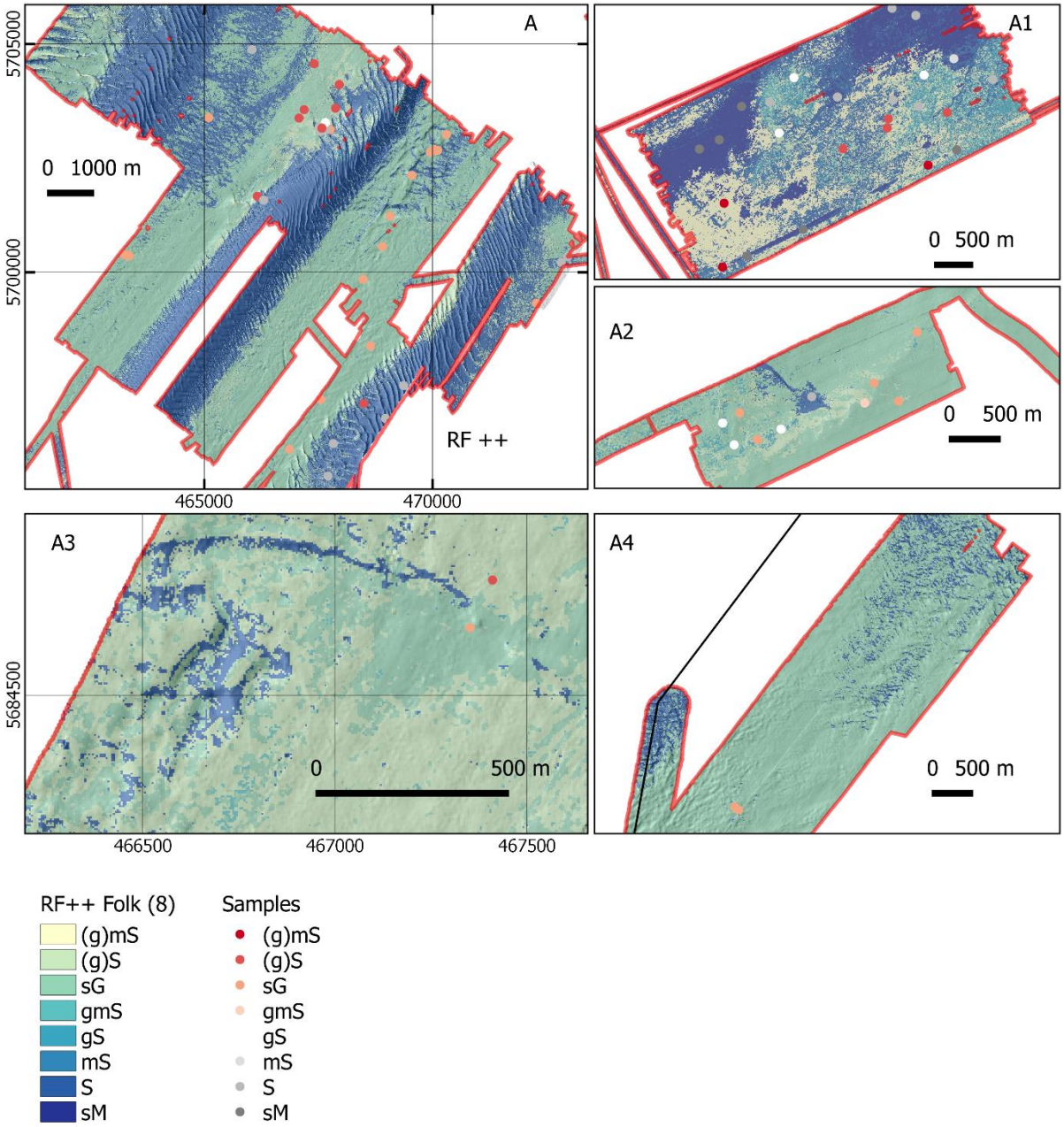


Figure 3.C1 – Maps produced by the RF++Folk (8) model. A) Hinder bank survey, A1) Oostende dumping site survey, A2) Thornton bank swale, A3) Buiten Ratel Rocky Zone, A4) North exploration area (ref. Fig. 3.1b for location of the areas in the BPNS).

Table 3.C1 - Raw confusion matrices for the RF++Folk (8) model classification. Accuracy metrics in the main text.

RF++ 8 Model Class	(g)mS	(g)S	sG	gmS	gS	mS	S	sM	Row Totals	User	Producer
(g)mS	**1**	1	1	0	0	0	0	0	3	0.33	0.33
(g)S	0	**2**	1	0	1	0	1	0	5	0.33	0.4
sG	0	1	**15**	0	0	0	1	1	18	0.79	0.83
gmS	0	0	0	**1**	0	1	0	0	2	0.5	0.5
gS	0	2	1	0	**5**	0	1	0	9	0.83	0.56
mS	0	0	0	0	0	**0**	1	1	2	0	0
S	1	0	1	1	0	1	**6**	2	12	0.55	0.5
sM	1	0	0	0	0	0	1	**1**	3	0.2	0.33
Column Totals	3	6	19	2	6	2	11	5	54		

Appendix D – Forward selection stepwise multiple regression analysis.

Here, following the methodological examples set out in Ferrini and Flood (2006), Multiple Regression Analysis (MLA) is used to determine relationships between backscatter (dependent variable) and sediment characteristics (independent variables), quantified by coefficients. The *lm* function in base R programming language was used to carry out the analysis. Hereafter, the results of the continued exploratory data analysis are presented. Given the at least partly semi-qualitative nature of the dataset (i.e. the gravel percentage being visually derived [i.e. grid count]), this analysis is not meant to capture fully quantitative relationships between the proposed variables and BS, rather, it provides more general insights into the joint drivers of the BS considering the entire study area and by considering survey areas in isolation. In this analysis, both continuous and categorical variables are used. Presence and absence of geogenic and biogenic material where tabulated for each ground truth location following inspection of the photographic material acquired on board for each sample. The justification for this analysis is the recognition of the non-univariate relationship between MBES BS and sediment type.

Treating the entire study area (i.e. the overall seamless backscatter dataset and overall number of ground truth samples), the multiple linear regression model with the highest statistical significance ($R^2 = 0.58$) was identified as:

$$dB = -35.7 - 4.3^{\text{No Gravel}} - 2.3^{\text{Yes Layering}} + 2.6^{\text{Yes Shell}} - 1.2^{\text{Yes Burrows}} + 0.007^{\text{D90}} + 0.005^{\text{Sorting}}$$

As such, it is found that it is the combined presence/absence of gravel (No Gravel; p-level: < 0.001), layering (Yes Layering; p-level: 0.001), bioclastic detritus (Yes Shell; p-level: < 0.001), bioturbation under the form of burrows (Yes Burrow; p-level: 0.1), sorting of the sediment grains in the sediment fraction 0.063-0.5 mm (Sorting; p-level: 0.01) and the larger percentage of coarse particles in the sediment fraction 0.063-0.5 mm (D90; p-level: < 0.001) which together explains the variance in the 300 kHz backscatter strength across the entire survey extent. Interestingly, the sign of the coefficients tells the direction of the identified trends, for example, either presence of coarse material such as gravel and or shell leads to significant increase in backscattering and oppositely, presence of either burrows and/or layering leads to decreasing backscattering.

The same modelling approach was applied to a set of sub-areas (i.e. those where 10 or more samples were available); namely: Kwinte Bank Swale, Thornton Bank Swale, the Oostende Dumping Site area, the Hinder Bank area (including swales and sand banks) and the Coastal trajectory area; the models are reported in Table 3.D1.

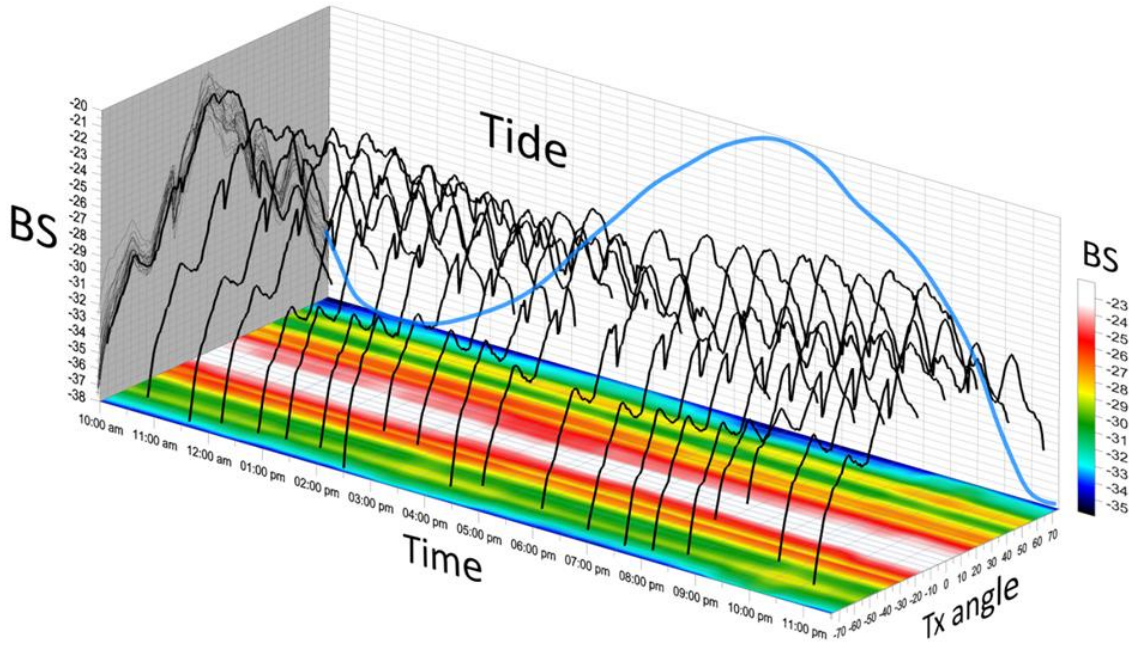
Table 3.D1 – Multiple linear regression models for each sub-area selected.

Area	Stepwise AIC R^2	Most significant model identified
Hinder Bank Area	0,55	dB = -35*** + 2,2 Pred Shell** - 2,6 Yes Burrow* - 0,01 D50(.) + 0,01 D90*** - 1,8 Skewness(.) + 4,3 Yes Gravel***
Coastal transect	0,56	dB = -39,5*** + 4.3 Pred Shell*** + 0,04 D50* - 0,07 D90* + 7.2 Skewness(.) + 0,01 Sorting*
Kwinte Bank Swale	0,93	dB = -102* + 0,09 Mode ** - 2,8 Yes Burrow* -0,35 D50** + 0,64 D10* + 25,3 D90/D10*
Thornton Swale	0,91	dB = -22** - 0,03 D50(.) + 0,05 D10* - 0,03 Sorting * + 5,9 Yes Gravel ***
Oostende dumping site	0,81	dB = -57*** - 0,16 D50* - 0,09 D10* + 0,06 D90** + Sorting 0,09** + Yes Gravel 2,45 (.)

Note: p-value code = *** 0; ** 0.001; * 0.01; (.) 0.05

Here, the assessed variables jointly explain the variance in 300 kHz backscatter in different ways and strengths at different survey areas. The associations found range from moderate ($R^2 \sim 0.5$) to very strong ($R^2 > 0.9$). however, for surveys taking place within swale morphologies (thus characterised by widespread presence of gravel and or shell clastic material), it is found that presence of gravel and shell is a statistically significant factor driving the backscatter response with p-levels consistently < 0.001. On the contrary, areas such as the Oostende dumping site, situated nearshore where relatively homogenous and moderately-sorted muddy and sandy sediments are predominant, it is a combination of grainsize parameters of the fine sediment fraction ≤ 0.5 mm (i.e. D10, D50, D90 and sorting) that drives the BS variance, leaving presence/absence of gravel and/or shell as a parameter of lesser importance.

Chapter 4



4. Insights into the short-term tidal variability of multibeam backscatter.

Field experiments on different seafloor types

Published as:

Montereale-Gavazzi, G., Roche, M., Degrendele, K., Lurton, X., Terseleer, N., Baeye, M., Francken, F., Van Lancker, V., 2019. Insights into the Short-Term Tidal Variability of Multibeam Backscatter from Field Experiments on Different Seafloor Types. Geosciences 9, 34. https://doi.org/10.3390/geosciences9010034

Presented and discussed at:

Conference: **Montereale-Gavazzi, G.,** Roche, M., Degrendele, K., Baeye, M., Francken, F., De Bisschop, J., Van Lancker, V., 2016. Can multibeam-derived acoustic backscatter be used to monitor changes in seabed habitats? North Sea Open Science Conference, 7-10/11/2016, Ostend (Belgium) (abstract + presentation)

Conference: **Montereale-Gavazzi, G.,** De Bisschop, J., Roche, M., Degrendele, K., Baeye, M., Francken, F., Van Lancker, V., 2017. In-situ appraisal of near-bed and water column particle transport on MBES backscatter. Geological and Biological Marine Habitat Mapping Conference, Halifax (Canada), 1-5/5/2017 (abstract + presentation)

Conference: Montereale-Gavazzi, G., Roche, M., Francken, F., Van Lancker, V., 2018. Short-term variability of high-frequency seafloor multibeam echosounder backscatter: results from field experiments on the Belgian Continental Shelf. GeoHab 2018. Marine Geological and Biological Habitat Mapping, Santa Barbara, Republic of California, U.S.A, 5-14/5/2018 (abstract + presentation + poster)

Invited Workshop: Mareano Infomar Maremap (MIM) Workshop - Hosted at British Geological Survey, The Lyell Centre, Edinburgh (Scotland) 27-28/02/2018. A field experiment investigating short term variability of MBES seafloor backscatter, (presentation).

Invited Talk. Backscatter Working Group piloted by the Geological and Biological Habitat Mapping Committee - 8/05/2018. Hosted at the Fess Parker Hotel Santa Barbara California, USA. Seafloor Backscatter Variability subgroup progress, (presentation).

4.1 Abstract

Three experiments were conducted in the Belgian Part of the North Sea to investigate short-term variation in seafloor backscatter strength (BS) obtained with multibeam echosounders (MBES). Measurements were acquired on predominantly gravelly (offshore) and sandy and muddy (nearshore) areas. Kongsberg EM3002 and EM2040 dual MBES were used to carry out repeated 300-kHz backscatter measurements over tidal cycles (~ 13h). Measurements were analysed in complement to an array of ground-truth variables on sediment and current nature and dynamics. Seafloor and water-column sampling was used, as well as benthic landers equipped with different oceanographic sensors. Both angular response (AR) and mosaicked BS were derived. Results point at the high stability of the seafloor BS in the gravelly area (< 0.5 dB variability at 45° incidence) and significant variability in the sandy and muddy areas with envelopes of variability > 2 dB and 4 dB at 45° respectively. The high-frequency backscatter sensitivity and short-term variability are interpreted and discussed in the light of the available ground-truth data for the three experiments. The envelopes of variability differed considerably between areas and were driven either by external sources (not related to the seafloor sediment), or by intrinsic seafloor properties (typically for dynamic nearshore areas) or by a combination of both. More specifically, within the gravelly areas with a clear water mass, seafloor BS measurements where unambiguous and related directly to the water-sediment interface. Within the sandy nearshore area, the BS was shown to be strongly affected by roughness polarization processes, particularly due to along and cross-shore current dynamics which were responsible for the geometric reorganization of the morpho-sedimentary features. In the muddy nearshore area, the BS fluctuation was jointly driven by high-concentrated mud suspension dynamics together with surficial substrate changes, as well as by water turbidity, increasing the transmission losses. Altogether, this shows that end-users and surveyors need to consider the complexity of the environment since its dynamics may have severe repercussions on the interpretation of BS maps and change-detection applications. Furthermore, the experimental observations revealed the sensitivity of high-frequency BS values to an array of specific configurations of the natural water-sediment interface which are of interest for monitoring applications elsewhere. This encourages the routine acquisition of different and concurrent environmental data together with MBES survey data. In view of promising advances in MBES absolute calibration, allowing more straightforward data comparison, further investigations of the drivers of BS variability and sensitivity are required.

Keywords: Multibeam echosounder, Seafloor, Backscatter, Monitoring, Short-term variability, Sensitivity, High-frequency.

4.2 Introduction

The North Sea is amongst the most highly impacted areas of the marine biome (Halpern et al., 2008). This is particularly the case for its Belgian Part where a multitude of anthropic activities, including intense routed navigation, dredging and disposal of dredged material, marine aggregate extraction, bottom trawling by commercial fisheries and extensive infrastructural, engineering and management developments (e.g. telecommunication pipelines, wind energy and beach nourishment) take place over a limited spatial extent of ~ 3600 km^2 along a ~ 65 km coastline (Douvere et al., 2007). In this regard, knowledge of the seafloor composition and of its spatio-temporal evolution is of great relevance to monitor human impacts on benthic habitats (of which substrate type is a fundamental abiotic component and surrogate for biota (McArthur et al., 2010)). At the European level, the monitoring is mandated by the European Marine Strategy Framework Directive to achieve Good Environmental Status (GES) of marine waters by 2020 (see (Rice et al., 2012) and references therein). Twelve GES descriptors were put forward for which each EU Member State defined indicators with associated monitoring programmes. For the Belgian Part of the North Sea (BPNS), one of them relates to changes in the extent of seabed habitats for which multibeam echosounding (MBES) was selected for the monitoring (Belgian State).

The use of MBES systems to acoustically characterize the seafloor has developed at a fast pace over the past three decades (De Moustier and Matsumoto, 1993; Brown et al., 2009). Co-registration of depth (signal travel-time) and reflectivity (backscattered intensity of the echo signals, hereafter BS) measured over a large range of angles (swathe) and at very-high resolution is possible using this technology. MBES BS depends on many factors including (Urick, 1967): (1) sediment type and its geotechnical characteristics dictating the seawater-seafloor impedance contrast (e.g. porosity, roughness, grain size and sediment inner homogeneity), (2) the sonar operating frequency, and (3) the signal angle of incidence. Due to the various sound-scattering properties of different seafloor substrates, BS can be used as a proxy aiding in the determination of bottom type at the water-sediment interface (e.g. Hughes-Clarke et al., 1996, Ferrini and Flood, 2006) and possibly the inference of some of its physical characteristics (Lamarche and Lurton, 2018; Lurton, 2010). Mapping this interface over vast areas allows extending information from local observations (in-situ ground-truth measurements) or transect-based information, that need interpolation/extrapolation (Strong et al., 2017), to the spatial continuum of the seafloor. This is valuable as an input to Marine Spatial Planning and Ecosystem Based Management and aids in the creation of efficient analytical, managerial and decision-making tools (Brown et al., 2012; Buhl-Mortensen et al., 2015, Madricardo et al., 2017).

Backscatter data obtained from MBES surveys are usually considered at two processing levels: angular response (AR - signal processing) and mosaicked images

(image-analysis). The AR describes the backscattering strength variation with angle of incidence and is retained as an intrinsic property of the seafloor directly relating to physical quantities of interest (Jackson et al., 1986). This “raw” format of backscatter is a promising seabed classification feature with a high potential for sediment discrimination as reported in a range of studies (Hughes_Clarge, 1997; Che Hasan et al., 2014 and references therein). The AR forms a shape (“the AR curve”) which reflects the dominant acoustic phenomena occurring along the angular domains: high-intensity specular reflection around nadir and lower-level scattering at oblique angles, strongly decreasing at shallow grazing angles. Where absolute calibration of the BS is achievable, the BS AR is to be considered as an objective measurement for which different methods exist (Eleftherakis et al., 2018 and references therein). The mosaic backscatter is a further derivative of the backscatter data, where BS levels are presented, usually in a georeferenced frame, in the form of a grayscale image with the angular dependence removed via statistical compensation. As such, the complete scene seems to be observed from the same incidence angle which is generally obtained by normalizing the data and referencing it to a conventional angle or a limited range of angles. Typically, this is around 45° where the angular dependence is weakest and where the sediment response dominates (Lurton, 2010). Both BS data forms (AR and mosaicked images) have been used to predict seafloor type, on their own, or in combination with other MBES data types (Che Hasan et al., 2014; Fonseca and Mayer, 2007). The main differences between these two formats are the spatial resolution and the type of information they contain. The BS AR is obtained by averaging a set of consecutive pings and processing them over the swath extent or over areas of interest, resulting in a resolution approximating that of the area selected. The BS mosaic resolution is considerably finer given it can be gridded as a function of the bathymetric resolution. Here, identification of small-scale features (down to decimetric orders of magnitude for high-frequency MBES operated in shallow waters) is feasible and is particularly valuable to ecological modelling requiring detailed discrimination of substrate distribution, down to the spatial-unit level of single patches (Galparsoro et al., 2009, Weding et al., 2011; Ierodiaconou et al., 2018). However, due to its inherent compensation of angle dependency, the mosaicking process leads to a loss of quantitative/physical information, making immediate ground truthing critical for effective relation to seabed properties. On the other hand, the AR can be interpreted via modelling of the response and fitting of parameters (see Lamarche et al., 2011 and references therein) which directly relate to the physical nature of the underlying substrate. Inversion of the AR into sedimentologically relevant information is a principle known for long which is currently hindered by a lack of high-frequency geoacoustic models dedicated to solving the” inversion problem” and should be perceived as an advancing application within the realm of acoustical oceanography. It however remains promising considering the rapid advances in MBES system absolute (Eleftherakis et al., 2018; Ladroit et al., 2018, Fezzani and Berger, 2018) and relative calibration (Weber et al., 2018) and in stability and repeatability controls (Roche et al., 2018), together promoting the comparability of data in space and time. It would also allow compiling

acoustic inventories that are calibrated against substrate types (and of associated features and combinations) to be used more globally. Alternatively, ground-truthing developments allow an increasingly detailed characterization of the acoustic observations and thus the potential development of models otherwise constrained by the need of a priori knowledge. On the long term, the scientific community would largely benefit from the development of detailed high-frequency geoacoustic models offering the advantage of directly exploiting the remotely-sensed data, thus reducing labour intensive and often expensive ground-truthing operations. Methods exploiting the AR demonstrated the utility of inverting radiometrically-calibrated and geometrically-corrected backscatter data into relevant sedimentological parameters (Fonseca and Mayer, 2007). However, the latter were related to well-sorted and homogeneous sediments only, evidencing the need to enhance the understanding of the relations between naturally-complex sediment configurations and the retrieved acoustic signatures and to ground truth the acoustics to avoid misleading interpretation (regardless the type of BS product and approach used).

Environmental monitoring, based on the acquisition of MBES time series (Roche et al., 2018; Lucieer et al., 2018), requires investigating and understanding the repeatability and variability of the data. Besides the instrumental constraints (aimed at ensuring the consistency of measured data from different campaigns and/or sensors), multiple sources of environmental factors must be considered for their impact on the consistency and accuracy of backscatter data measurement. This is particularly the case in nearshore/coastal and continental shelf zones where seafloor and water-column variability may be high at diverse scales in space and time. Therefore, it is important to evaluate whether changes in the average backscatter level between different surveys reflects actual changes in sediment properties or in the conditions of the water medium (Roche et al., 2015) and of other dynamic parameters. A similar concern was already identified in terrestrial remote-sensing applications (Singh et al., 1989; Floriciou and Rott, 2001). In this regard, it is critical that the survey-design phase of any such investigation considers all possible sources of variation which may contribute to unwanted fluctuations of the backscatter strength. This is needed to confidently quantify seafloor type and change based on the acoustic returns.

Depending on the MBES survey environment, a range of factors can be responsible for unwanted signal fluctuations in the acoustic measurements. First, the azimuthal dependence is driven by the orientation of small-scale bed forms relative to the navigation heading (hence the acoustic line of sight; see Ferrini and Flood, 2006; Boheme et al., 1984; Briggs et al., 2001; Richardson et al., 2001; Lurton et al., 2018)), as well as by seafloor mobility under the effect of hydrodynamic forcing driving the roughness polarization. Second, the dissipative nature of the water medium leads to absorption of acoustic energy during the signal propagation: it depends on the seafloor-target range, frequency and physico-chemical properties such as temperature and salinity driving the viscous-thermal status (Francois and Garriosn, 1982a, b; de Campos Carvalho et al., 2013; Richards et al., 1996). The

concentration and particle size of suspended particulate matter (SPM) also contribute to the total two-way transmission loss of the acoustic signal; it can be significant in nearshore and shelf environments (particularly over relatively long distances, i.e., typically beyond 100 m – see Richard et al., 1996). Finally, biological activity, occurring in the water column (e.g. the Deep Scattering Layer – see Holliday et al., 1998) or at the benthic level (referring to epibenthic and infaunal activity – see Briggs et al., 2001; Gorska et al., 2018), as well as near-bed advection of submerged aquatic vegetation (Madricardo et al., 2017) can affect MBES measurements. Additionally, there is a need to better understand the effects of the intrinsic dynamicity of given substrates and how near-bed (also referred to as boundary and/or water-sediment interface and benthic zone) sediment transport affects the seafloor sonar detection. Ideally, all of these variables are accounted for when comparing datasets in space and time.

This study presents a set of observations originating from three experimental datasets acquired to understand and quantify the external and seafloor-intrinsic sources of variance that may lead, while surveying, to biases in the seafloor backscatter acquired by high-frequency (300 kHz) multibeam sonar systems. Repeated measurements (multi-pass MBES surveys) using EM3002D and EM2040D echosounders are interpreted based on seafloor and water-column data acquired by grab sampling, optical observations and a multi-sensor benthic lander, in combination with a drop-down frame. Altogether, these data are used to assess the sensitivity of the BS and how its short-term variability can affect the detection of actual changes in the seabed.

4.3 Materials and Methods

4.3.1 Description of MBES and survey areas

Multibeam data were collected using Kongsberg EM3002D and EM2040D echosounders, respectively installed on RV A962 *Belgica* (http://odnature.naturalsciences.be/*Belgica*) and RV *Simon Stevin* (http://www.vliz.be/en/rv-simon-stevin). Table 4.1 reports the parameters used to operate the echosounders during the experiments.

Table 4.1. MBES specifications and main settings, and associated ancillary sensors

Parameter/Echosounder	Kongsberg Maritime EM3002D	Kongsberg Maritime EM2040D
Number of soundings per ping	508	800
Central frequency	300 kHz	300 kHz
Pulse length	150 µs	108 µs
MBES Mode	Normal	Normal
Rx Beam spacing	High density equidistant	High density equidistant
Tx x Rx Beam width	1.5° x 1.5°	1° x 1°
Positioning System	MGB Tech with Septentrio AsteRx2eH RTK heading receiver	MGB Tech with Septentrio AsteRx2eL RTK receiver
Motion Sensor	Seatex MRU 5	XBlue Octans
Sound Velocity Probe	Valeport mini SVS and SVP	Valeport mini SVS and SVP

Three surveys were conducted during spring-tide regime: in February 2015, March 2016 and November 2017, respectively on the Kwinte swale, Westdiep swale and MOW 1 areas featuring distinct seafloor substrates. Locations are displayed in Figure 4.1 and general environmental conditions are given in Table 4.2. Within the areas, study sites were selected with homogeneous acoustic signatures, based on previous surveys and ancillary data.

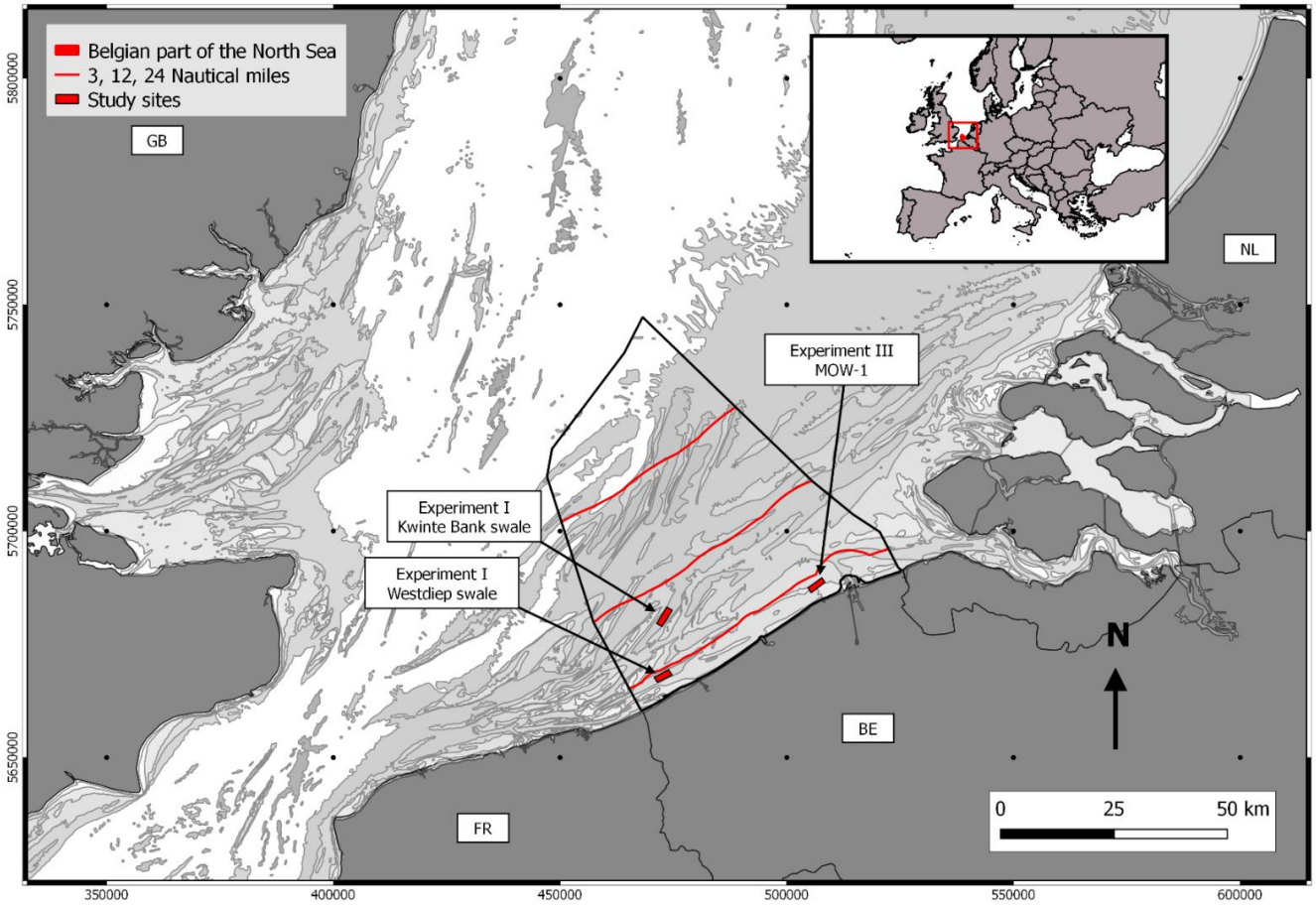


Figure 4.1 - Location of selected study sites within the Belgian Part of the North Sea: (1) Kwinte swale area (central coordinate: N 51° 17.2717, E 002° 37.7035), (2) Westdiep swale area (N 51° 09.1230, E 002° 34.6806), (3) Zeebrugge, MOW 1 pile area (N 51° 21.6697, E 003° 06.5798). Data are projected in World Geodetic System 84 (WGS 84) in Universal Transverse Mercator Zone 31 N (UTM – 31N). This coordinate system is used throughout the rest of the document.

Table 4.2 - Environmental characteristics of the three experimental areas, each having distinct seafloor substrate properties. MLLWS: Mean Lowest Low Water at Spring tide.

Area	Depth and sediment dynamics*	Habitat type (EUNIS level 3**)	Details on Environmental setting
Kwinte swale	Depth (MLLWS): 25 m Water mass type: clear seawater Magnitude of sediment transport during Spring tide: < 0.5 tonnes $m^{-1}d^{-1}$	Offshore circalittoral gravelly hummocky/hillocky terrain (relatively well sorted medium sand with gravel)	In Roche et al. (2018) and Bellec et al. 2010)
Westdiep swale	Depth (MLLWS): 15 m Water mass type: clear seawater Magnitude of sediment transport during Spring tide: 0.5-1 tonnes $m^{-1}d^{-1}$	Circalittoral sandy/siliciclastic terrain (well sorted fine to medium sand)	In Van Lancker et al. (1999)
Zeebrugge, MOW1 pile	Depth (MLLWS): 10 m Water mass: Turbidity maximum zone Magnitude of sediment transport during Spring tide: > 1 tonnes $m^{-1}d^{-1}$	Circalittoral muddy sediments	In Fettweis and Baeye (2015) and Baeye et al. (2012)

*From Lancneus et al. (2001) and Van Lancker et al. (2007)

** European Nature Information System level III categories – see Davies et al. (2004)

4.3.2 Survey methodology and data processing

The surveying principle designed to capture short-term backscatter variability over the same seafloor patch is presented in Figure 4.2. It consists of a series of repetitive MBES measurements performed over the duration of a tidal cycle (~13h). The same reference survey-line (~ 2 km) was followed using the same heading and crossing the centre of a region of interest (ROI – approximately 500 x 200 m for the first two experiments and 200 x 50 m in the third one). While deviations from the planned track line could happen for several reasons, this did not occur significantly during the experiments and the homogeneity of the selected ROIs ensures the spatio-temporal consistency of the data across all insonified angles. Runtime acquisition parameters used in the Kongsberg Seafloor Information System software suite (Kongsberg Maritime) were kept rigorously unchanged throughout the duration of each experiment avoiding introducing extra sources of variance in the data.

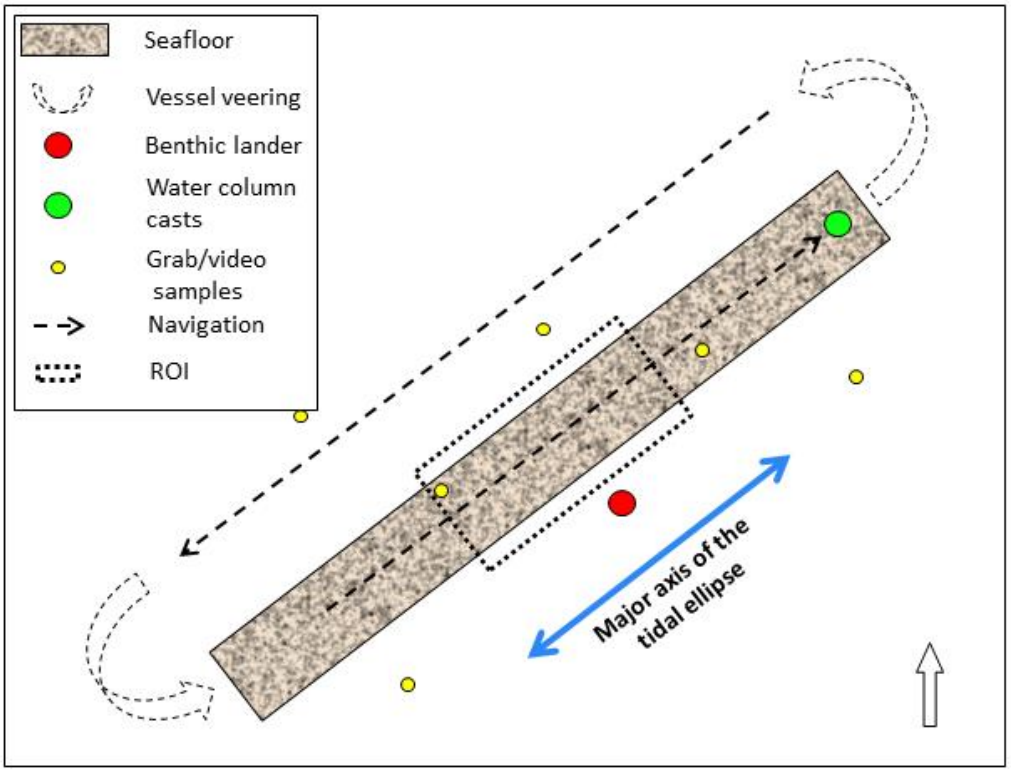


Figure 4.2 - Schematic representation (not to scale) of the surveying principle designed to capture the short-term backscatter variability over a homogeneous region of interest (ROI). See main text for explanations.

Each experiment consists in the acquisition of a short-term backscatter and bathymetry time series according to the described strategy. To interpret the acoustic data, different strategies were put forward to quantify environmental variables during the experiments; these are listed hereafter for each experiment.

4.3.2.1 Experiment 1 – Kwinte swale area

The first experiment alternated MBES measurements with vertical profiling of oceanographic variables using a drop-down frame over a 13-h tidal cycle. The area was selected because of its high stability in MBES-measured BS, based on previous investigations. Meanwhile, this site was proposed as a natural reference area to control the BS stability prior to any surveying operation in the Belgian Part of the North Sea (BPNS) (Roche et al., 2018). The oceanographic data relating to this experiment are discussed in Roche et al., (2015) and De Bisschop (2016). They show negligible effects of water-column processes and of near-bed sediment transport on the backscatter measurements. Here, only the MBES data are discussed.

4.3.2.2 Experiment 2 – Westdiep Swale area

The second experiment was extended with the deployment of a benthic lander equipped with oceanographic sensors (Figure 4.3; Table 4.3) from which variables relating to the lower ~ 2.4 m above seabed (mab) were derived. The lander was moored at ~120 m distance from the nadir of the MBES track line. This was the minimum distance allowed to keep a safe navigation buffer from the instrument's signalling buoy. Given the similar morpho-sedimentary characteristics over the survey area, the information sampled by the lander was considered as representative of the processes occurring within the MBES ROI.

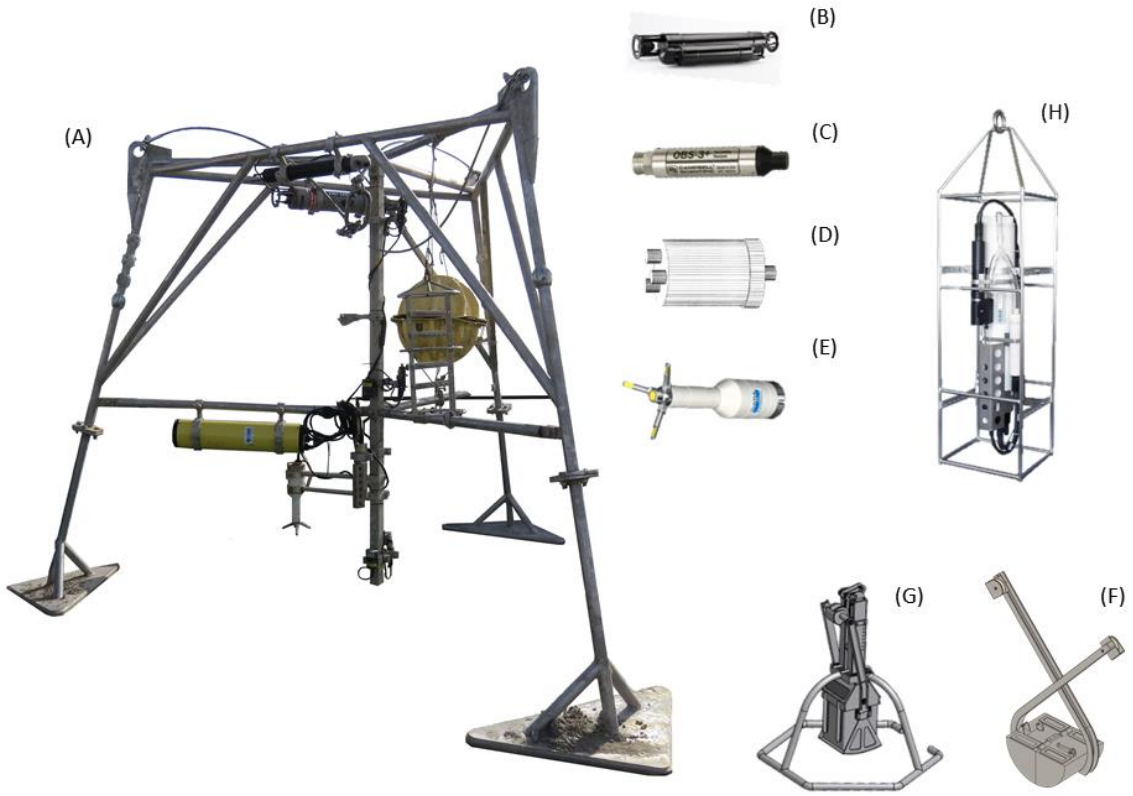


Figure 4.3 - A) Benthic lander equipped with a set of oceanographic sensors (see Table 1 for details about the instrumentation) deployed during the second experiment in the Westdiep study site. A similar lander was deployed for the third experiment. A chain of OBS+ sensors at 0.3, 1 and 2.4 meters above bottom (mab) was present during the third experiment. In this image: A) Benthic lander frame, B) Laser In-Situ Transmissometer, C) Optical Backscatter Sensor, D) Acoustic Backscatter Sensor, E) Acoustic Doppler Velocimeter; and on-board winch-operated instruments, F) Van Veen Grab, G) Reineck box core, H) CTD frame, equipped with a OBS+ and a Niskin bottle.

Measurements of suspended particulate matter concentration (SPMc) were derived using optical and acoustic backscattering sensors (OBS and ABS). Field calibrations of the OBS were carried out during previous RV *Belgica* cruises following the methodology described in Fettweis and Baeye (2015). Despite the calibration locations being different, derived SPMc are sufficiently representative of water-column processes occurring at 2.35 mab in the current study area. The multi-

frequency ABS was equally used to determine SPMc, as well as median grain size (D50) in a 1-m profile above the bed and per bins of 1 cm. This sensor was chosen due to its suitability to measure in sandy environments. Calibration is provided by the manufacturer (implicit calibration methods; see Thorne and Hanes, 2002), and is based on the use of glass spheres being representative of quartz/siliciclastic particles present in this study area. Along with MBES and benthic lander data, an SBE 19+ SeaCAT Profiler CTD, equipped with a 5L Niskin bottle, was regularly down casted at the end of the MBES transect to obtain measures of SPMc, salinity, depth and temperature in the water column up until ~3 mab. This was performed approximately every hour. From each water sample, three sub-samples were filtered on board using pre-weighed filters (Whatman GF/C type). In turn, they were subsequently washed with 50 ml of Milli-Q water to remove salt, dried and weighted to derive SPMc. MBES and all benthic lander data were referenced to a uniform timestamp (the mean time of acquisition within the defined ROI) to enable later inter-comparison.

Additionally, a set of reconnaissance Van Veen grab samples (n = 7, replicate = 3) were acquired in the surroundings of the experiment site and were analysed for grain size by means of a Malvern Master-sizer 3000 (www.malvern.com). Before the analysis, organic matter and calcium carbonate ($CaCO_3$) were removed using H_2O_2 (35 %) and HCl (10 %), respectively. To describe sediment types, the Folk and Ward nomenclature available from the GRADISTAT (Blott and Pye, 2001) software is used throughout the rest of the document.

Table 4.3 - Summary of the oceanographic sensors installed on the benthic lander used to quantify the driving processes of variability in the MBES backscatter measurements.

Sensor	Measurements/variables	Distance of measurement from seabed	Temporal/spatial resolution	Further instrument specifications	Calibration
ADV Ocean velocimetry @ 5 MHz	Current in x, y, z; Direction; Altimetry; Temperature; Salinity; Velocity	0.2 mab	Bursts of 15 min. 2x2 cm measuring cell	www.sontek.com	NA
ABS Acoustic Backscatter Sensor @ 0.5, 1, 2, 4 MHz	SPMc; particle size	1 mab	Bursts of 30 min. 1 cm bins over 1 m profile	www.aquatecgroup.com	Manufacturer calibration (implicit method)
Sequoia Scientific LISST 100-X (type-C)	Particle size and distribution; transmission; volume concentration	2.4 mab	Bursts of 1 min.	www.sequoiasci.com	NA
OBS+	SPMc	2.35 mab	Bursts of 15 min.	www.campbellsci.com/d-a-instruments	Previous campaign calibration using in-situ water samples (gravimetric analysis)
SBE 19+ SeaCAT Profiler CTD – OBS+ and 5L Niskin bottle	Temperature, Salinity, hydrostatic pressure; SPMc (from water filtrations of Niskin bottles)	~ 2/3 mab	~ Every 1 h	www.campbellsci.com/d-a-instruments and www.seabird.com	OBS NTU* vs SPMc Calibration = R^2 0.56 @ 3 ~ mab

*NTU: Nepheloid turbidity units

4.3.2.3 Experiment 3 – Zeebrugge, MOW 1 Pile area

The third experiment was carried out in the proximity of a fixed monitoring station (MOW 1 - http://departement-mow.vlaanderen.be) where a benthic lander is deployed regularly by the Royal Belgian Institute of Natural Sciences as part of a long-term sediment dynamics monitoring programme (Fettweis and Baeye, 2015). The benthic lander allowed obtaining SPMc from a set of turbidity meters installed at 0.3, 1 and 2.4 mab. The OBS signals have been related to mass concentration after

calibration using mass-filtered water samples, taken during a 13-h tide cycle. Furthermore, during this experiment, a time series of Reineck box cores was also collected to quantify changes in surficial sediment composition over the duration of the experiment. Overall, 12 samples were collected (approximately one every hour). They were taken from a relatively homogeneous seafloor patch and within a buffer zone with a radius of ~100 m. Particle sizes were analysed, and their nature was described as specified in the previous section. To obtain data relating to the immediate seabed surface of the samples, a 1-cm slicing was carried out on-board; the first three centimetres were kept for analysis. Additionally, two full-coverage surveys (covering approximately 350 m x 1.5 km) were acquired over this study site on November 21st and 24th of 2017 (experiment taking place before the second survey on the 24th November). Similarly to the acquisition of the time-series datasets, surveys were conducted by maintaining fixed runtime parameters and following the same set of navigation lines. Furthermore, both surveys were carried out during the same tide-window: around peak ebb flow. Following a routine to objectively find the statistical number of classes in the datasets (i.e. Within Group Sum of Squared Distances plot), maps were classified using the unsupervised k-means clustering algorithm (Hartigan and Wong, 1979) and assessed for changes by means of simple algebraic change detection (i.e. image differencing). This was carried out to appraise the short-term spatial sediment dynamics of the study area.
Considering the muddy and soft nature of the water-sediment interface of this study site and the chance to have ephemeral deposition of unconsolidated sediments Baeye and Fettweis, 2012), the Kongsberg Quality Factor (QF) was computed within the ROI to assist in the interpretation of the BS temporal/tidal oscillation. The QF is a metric relating to the relative bathymetry uncertainty and is expressed by the ratio between the scaled standard deviation of the range detection divided by the detected range (Kongsberg Maritime): the smaller the QF values, the smaller the uncertainty, implying a more accurate bottom detection. In this instance, the QF can be interpreted as a proxy of changes in the water-sediment interface, and thus for variability/sensitivity of the BS. Values of SPMc and QF are later related to the MBES BS time series by means of correlation and regression analysis.

4.4 MBES Processing

Different BS products were derived from the Kongsberg datagrams by using different software tools. All BS data were taken within the selected ROIs. Similarly to the acquisition phase, a rigorous standardized processing procedure was maintained to avoid variability induced by changes in software parameters (Roche et al., 2018). Using the QPS FMGT© module (QPS), time series of 1-m horizontal resolution mosaicked backscatter were produced. The default FMGT Geocoder compensation algorithm compensates the data over the angular interval from 30° to 60°. Secondly, using the SonarScope© software suite (SonarScope), time series of AR curves were derived from the Beam intensity datagrams. The seafloor angular backscatter

strength is computed from the following sonar equation liking the transmitted and received signal levels with the transmission losses and the backscattering process:

$$\mathrm{EL}\,(\mathrm{R},\Theta)\;=\;\mathrm{SL}-\;2\mathrm{TL}\,(\mathrm{R})\;+\;10\log\mathrm{A}\,(\mathrm{R},\Theta)\;+\;\mathrm{BS}\,(\Theta),$$

where EL is the Echo Level (referenced to 1 µPa) measured at the receiver as a function of the sonar-to-target range R and the angle of incidence Θ of the signal onto the seafloor, SL is the Source Level (in dB re 1 µPa @ 1 m), $2\mathrm{TL}$ is the two-way Transmission Loss accounting for both geometrical spherical spreading (i.e. $40\log\mathrm{R}$) and absorption ($2\alpha_{\mathrm{w}}\,\mathrm{R}$ – see Francois and Garrison 1982a, b), A is the instantaneously insonified area, delimited by the MBES beam aperture and/or signal duration, and BS is the Backscatter Strength of the seafloor target at the observation angle Θ. The data reduction scheme relating to the AR data-type is reported in Table 4.4 and, despite being relative, is considered to be the best estimate of the raw BS angular response (Fezzani and Berger 2018; Roche et al., 2018). Figure 4.4 shows the differences between AR prior and after removing the Kongsberg built-in Lambertian and specular adaptive corrections (the latter is removed a priori in the SonarScope® processing workflow). Time series of bathymetry for each experiment were also derived using QPS QIMERA© (Qimera). Tidal corrections using data from the closest tide-gauges were applied for the EM3002D datasets whereas a higher accuracy RTK (Real Time Kinematic) correction was applied to the EM2040D data.

The bathymetric time series were needed to assess morphological changes from 2D-depth profiles and 3D visualisation (for example between ebb and flood tidal phases). The vertical accuracy (at a 95 % confidence level from descriptive statistics of the conducted measurements) of the EM3002D is ± 4 cm, similarly to the value reported in Ernsten et al., (2006) and compliant with the accuracy obtained by the Continental Shelf Service of Belgium conducting periodically repeated measurements over a lock situated in the harbour of Zeebrugge and where the absolute depth is known. The vertical accuracy for the EM2040D data is yet not determined. Its IHO confidence interval (IHO, 2008) is around ± 15 cm, being too large to account for decimetric vertical changes. A 1-m pixel horizontal resolution was chosen as a good balance between the size of the insonified area at nadir and that insonified at shallow grazing angles.

Table 4.4 - Backscatter processing steps after Roche et al., (2018) for the AR time-series dataset (SonarScope© processing).

1.	Correction for sound absorption based on surface seawater properties (from the RV *Belgica* On-board Data Acquisition System - https://odnature.naturalsciences.be/*Belgica*/en/odas)
2.	Correction of the instantaneous insonified area using the real incidence angle as from the tide-corrected terrain model of the study site: the bathymetric surfaces are used to correctly allocate the backscatter snippet traces from single pings to their true seabed position.
3.	Removal of all angle-dependent corrections introduced by the manufacturer (e.g. the Lambert and specular corrections in Kongsberg Maritime MBES data).
4.	Per ROI: Computation of AR curves.

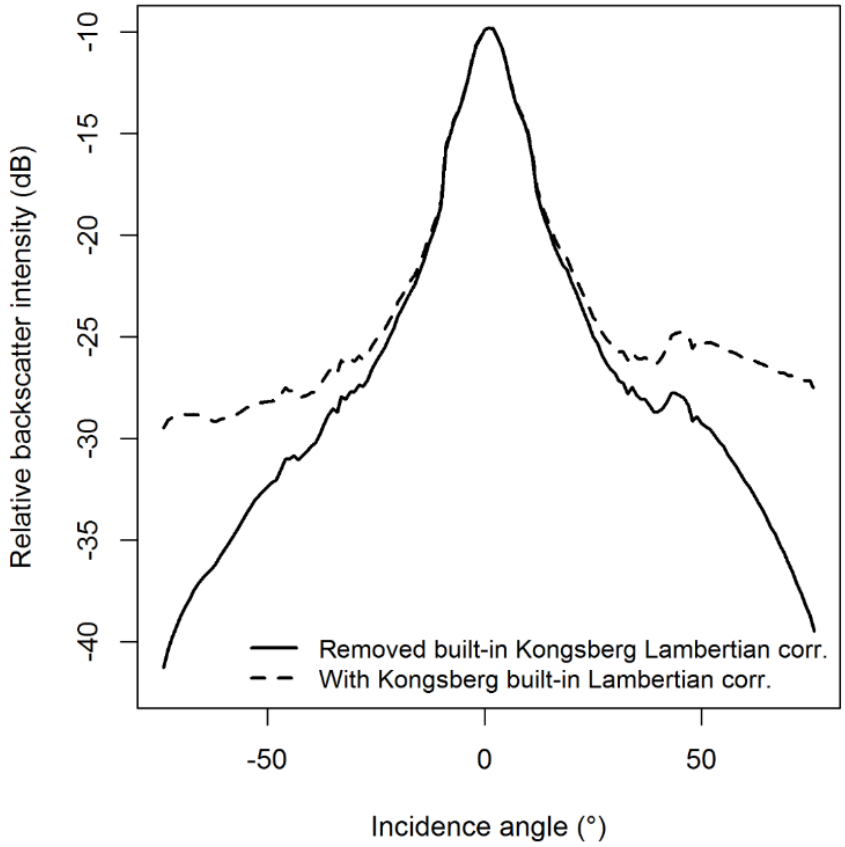


Figure 4.4 - Illustration of the difference between angular response curves provided by the Kongsberg manufacturer after correction in SonarScope© to remove the specular correction (dashed line, applied by default in SonarScope© processing routine) and the Lambertian correction (solid line, backscatter status 1 in SonarScope©). The solid line is the type of angular response data used in the present investigation and is believed to be the best estimate of the raw intrinsic seafloor backscatter response. The type of BS data output particularly suits the study of variability (i.e. relying on an artefact- and bias-free dataset) since the built-in specular-adaptive and Lambertian corrections are computed on a ping-to-ping basis, hence possibly introducing biases due to the local seafloor configuration.

4.5 Transmission losses

Different mechanisms beyond the inherent geometrical (spherical) spreading of the sound wave control the attenuation during the propagation in seawater and can be responsible for unwanted signal fluctuations and degradation of the signal-to-noise ratio (Lurton, 2010). Retrieval of the correct target backscatter strength must account on the dissipative nature of the seawater medium absorbing part of the acoustic energy via chemical reactions, viscosity and scattering (Lurton, 2010). Overall, attenuation losses (i.e. accounted by empirically-derived absorption coefficients within the $2\mathrm{TL}$ term of the sonar equation) result from the contributions of: (1) absorption in clear seawater (α_w) sensu Francois and Garrison (1982a, b) and (2) viscous absorption (α_v, Urick, 1948) and (3) scattering due to the presence of suspended particulate matter (α_s, Richards et al., 1996 ; Hoitink and Hoekstra, 2005).

The uncertainty introduced by the attenuation of sound (in dB/km) in seawater only was estimated for each experiment for nadir (0°), oblique (45°) and fall-off angular regions (70°). For the second experiment, the absorption model by Francois and Garrison (1982a, b) was applied to the set of water-column profiles ($\mathrm{n} = 10$) obtained by the CTD frame down-casts; for the two other experiments, only surface values of absorption coefficient were considered.

Using the modelling approach by Richards et al., (1996) and Hoitink and Hoekstra, (2005), sound absorption due to presence of suspended sediment (that due to combined viscosity and scattering) was estimated for the second and third experiments based on the available data (the routine was implemented in MATLAB ©). For the second experiment, this uncertainty was estimated for the 1- m profile above seafloor using the vertically-averaged ABS-derived SPMc and median particle size (D50) for the duration of the experiment.

Additionally, uncertainty was estimated along the quasi-continuous sediment profile (~15 m depth) that was reconstructed combining observations from the various sensors (i.e.; filtrations from the Niskin samples and the benthic lander mounted OBS and ABS sensors). The profile was reconstructed, and assumptions were made to represent a worst-case scenario, thereby selecting the data from the moments of maximal volume concentration. As such, the profile relates to 0.05 g/l from surface to 3 mab, 0.1 g/l from 3 to 0.5 mab and 0.3 g/l from 0.5 mab to seafloor. To appraise the effect of particle size, the D50 of the lower part of the profile was altered from 100 to 400 µm (reflecting the sand particles potentially resuspended in the near-bed of this area during spring tide). Despite a lack of data to carry out a similar analysis in the third experiment, the available OBS-derived SPMc time series were coupled to the MBES BS by means of correlation analysis and further descriptive plots to observe relationships. Nonetheless, similarly to the second experiment, the effect over the full water depth was estimated by reconstructing a quasi-continuous sediment profile based on values of volume concentration from the OBS chain and

using a fixed D50 of 63 µm (representative of suspended mud particles, characterising the turbidity of this area). Peak concentration values were selected here too, leading to a reconstructed profile of 0.2 g/l from surface to 2.5 mab, 1 g/l from 2.5 to 0.5 mab and 2 g/l for the lowest 0.5 mab. The effect of particle size was investigated here too, changing the D50 of the lowest part of the profile from 63 to 125 µm (approximating to the fine sand observed in the grab samples). For both cases, the transmission losses due to this factor are presented for nadir (0°), oblique (45°) and fall-off angular regions (70°) and for the described profile arrangements.

4.6 Results

4.6.1 Results display

This section presents the results of the three experiments. First, the spatial context is provided through gridded backscatter and bathymetry data products (Figure 4.5). Next, a synthesis is given on the short-term variability in the backscatter time series. Interpretation of the results is helped by the ground-truth data collected for experiments II and III: for the second experiment, the benthic lander data were summarized and used to produce a set of correlations between backscatter and variables; for the third experiment, interpretation of the BS spatio-temporal behaviour is supported by a Reineck-box core time series, the SPMc obtained by the OBS chain (n = 3, at: 0.3, 1 and 2.4 mab) on the benthic lander, the bathymetric uncertainty metrics and the full-coverage surveys acquired. For each experiment, results relating to the transmission losses are presented in a separate section.

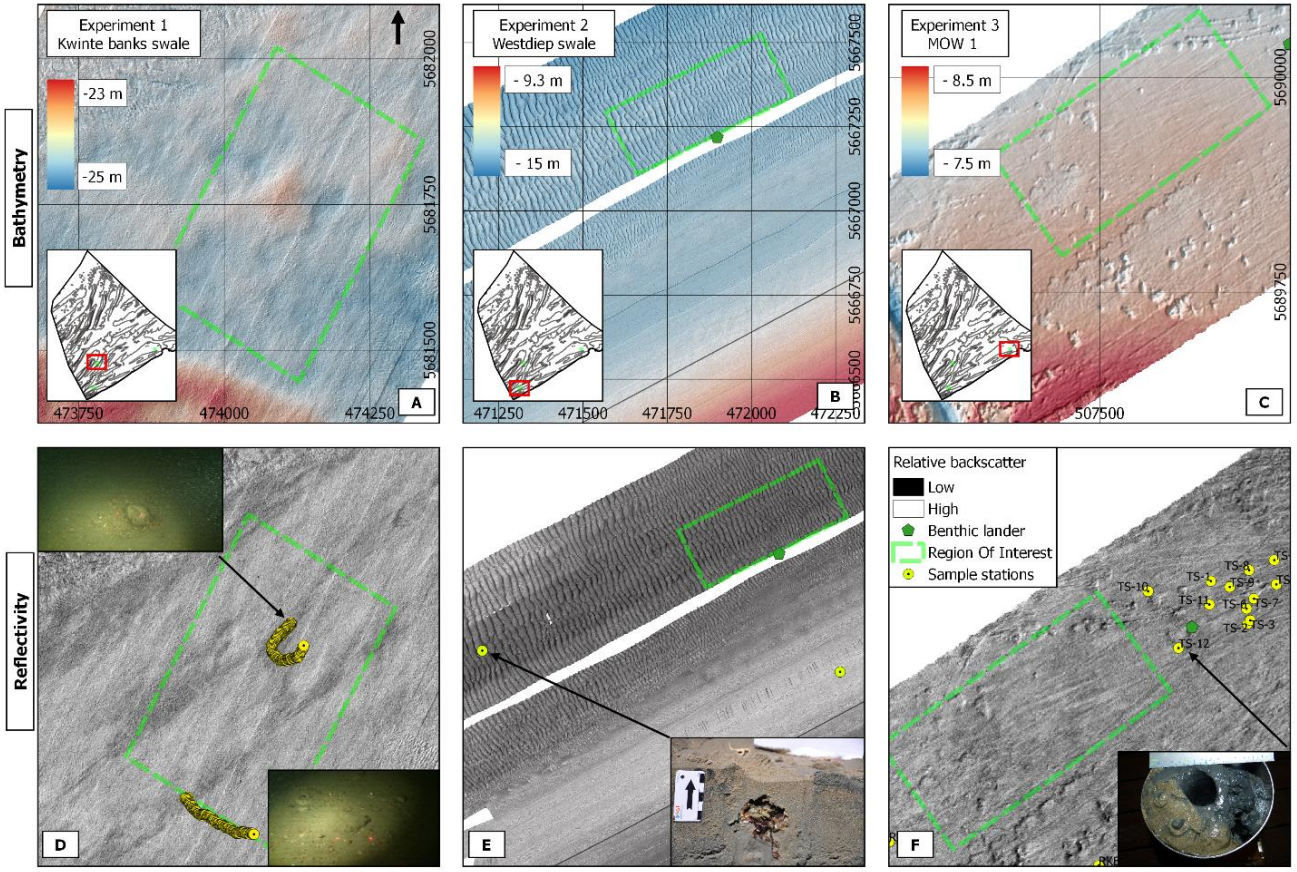


Figure 4.5 - Details of the bathymetry (A-B-C) and reflectivity maps (D-E-F) for each study area. For experiments II (B-E) and III (C-F), the location of the benthic lander, equipped with various oceanographic sensors, is denoted by a dark-green pentagon. Ground-truth stations are denoted by yellow circles, whereas the ROIs are denoted by green dashed-line polygons. Photographic details of the substrate types are also shown: for the Kwinte swale area images (D), the laser points are 9 cm apart (Courtesy of A. Norro, Royal Belgian Institute of Natural Sciences). Severe modification of the seabed by bottom trawling gears is noticeable at the MOW 1 study site (C, F): patterns of substrate erosion (elliptical depressions of ~ 10 to 30 cm in depth and up to 15 m in diameter) occur in the immediate proximity of the trawl marks.

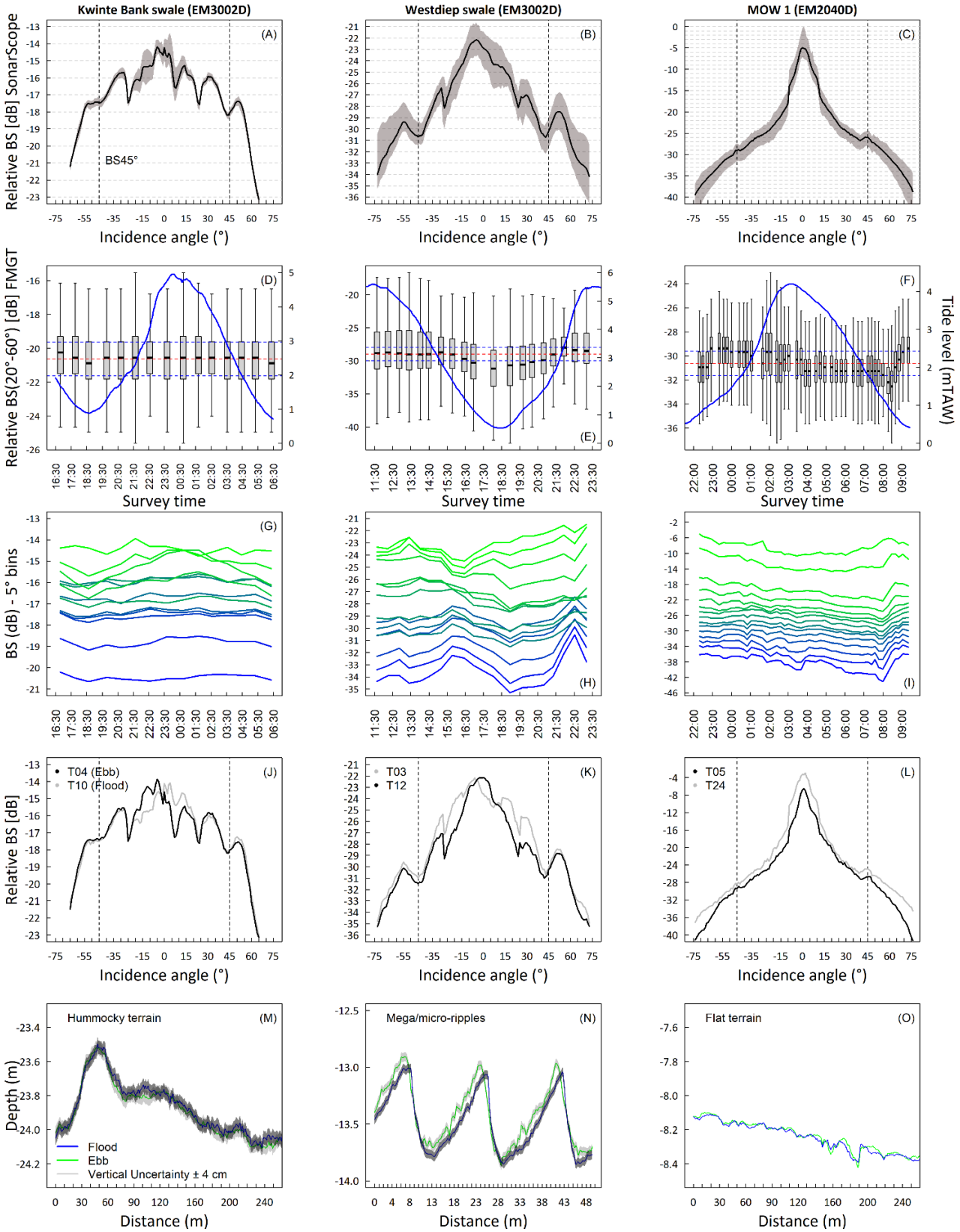


Figure 4.6 - Synthesis of the backscatter time series acquired for each experiment (continues on next page).

Figure 4.6 - Synthesis of the backscatter time series acquired for each experiment. The first plot (A-C) is the envelope of variability (grey shading) around the average AR (black line) of the full AR BS time series, extracted from the defined ROIs. It describes the variability of backscatter intensity per angle of incidence over the duration of the experiments. The envelope is computed from n = 15, 19 and 47 MBES passes respectively for the 1st, 2nd and 3rd experiment. The processing scheme code for the AR BS dataset is "A4 B1, C2 D1 E5 F3 G2 H3 I0 J0 H2" using the nomenclature proposed in Lamarche and Lurton (2018). The second plot (D-F) is the same time series (though derived from the BS mosaics produced in FMGT; $BS_{30\text{-}60° @ 300\ kHz}$) but visualized as boxplots of relative BS (values across the full incidence angle) against the time of acquisition (mean surveying time within the ROI). The overall mean over the full time series, together with the ± 1 dB Kongsberg sensitivity threshold (Hammerstad, 2000), are respectively shown as red and blue dashed lines. The tidal level is superimposed to assess a prospective BS trend in respect to the tidal oscillation and its phases. In the boxplots, lower and upper box boundaries are the 25th and 75th percentile respectively, the black central bar the median, whiskers denote the full extent of the data (i.e. min/max). The processing scheme code for the mosaicked BS dataset is "A4 B0 C0 D0 E5 F0" using the nomenclature proposed in Lamarche and Lurton (2018). The third plot (G-H) is the time evolution of the relative BS for areas insonified within a same envelope of incidence angle at a 5° resolution. This provides a more detailed depiction of the variability as a function of the incidence angle, to observe if smaller angular sectors would be less affected by the processes driving the variability. In G-H, the blue to green palette represents angular intervals from the fall-off to the specular region in steps of 5°, leading to approximately 15 sub-sectors per experiment. The fourth plot (J-L) displays the AR curves at the peak flood and ebb tidal phases (the legend mentions the corresponding survey time) during the experiments and is used to establish the presence of roughness-polarization dependence (as proposed in Lurton et al., 2018). The fifth plot (M-O) displays bathymetric profiles extracted at nadir within the ROIs at the same peak flood and ebb tidal moments as the previous plot (J-L). For the Kwinte swale and Westdiep experiments, using the EM3002D echosounder, the ± 4 cm vertical accuracy interval is displayed as a grey/transparent envelope.

4.6.1.1 Offshore gravelly area - Kwinte swale

Figure 4.5A and 4.5D show respectively details of the bathymetry and the backscatter for the Kwinte swale area. Sampling stations are also shown in this image (yellow circles). The sediment of this area is medium sand with gravel and bioclastic detritus and the seafloor presents a hummocky terrain typical of predominantly gravelly and shelly substrates of gullies (thalwegs) found in between the sandbanks of the BPNS. These substrate features were observable from the video imagery to be homogeneously distributed (with sporadic occurrence of boulders). The backscatter image for this area (Fig. 4.5D) is moderately uniform and presents a relatively high reflectivity throughout. The pattern observable relates to the tidal-ellipse orientation (SW-NE) which follows the main axis of the gully within which the site is situated (Van den Eynde et al., 2010).

The results of the repeated MBES data acquisition in this area are shown in Figure 4.6 (first column). The AR and boxplot time-series plots (Fig. 4.6A and 4.6D) denote the high stability of the sediment backscatter in the area over the duration of the tidal cycle. No trend is detectable. The interquartile range is about 2 dB, indicating a high homogeneity. The consistency of the time series (Fig. 4.6A and 4.6G) indicates that the short-term backscatter variability remains < 0.5 dB across all incidence angles,

except for the specular angular region (0°-18°) where the backscatter variability reaches up to ~2 dB. This behaviour is likely explained by a dependence related to the oscillations of micro-ripples (polarization under hydrodynamic forcing) which are beyond the imaging capability of the MBES spatial resolution. Figure 4.6J illustrates this behaviour as the AR curves at peak ebb and flood diverge more importantly in the specular angular region but converge above 25°. Interestingly, since the variability in the specular region is limited to an angle around 18°, it does not affect the mosaic production in FMGT Geocoder engine which compensates the data based on an angular interval ranging from 30° to 60°. Small depth differences (Fig. 4.6M) remain within the vertical accuracy of the soundings with only slight differences in profile indentation: this is likely indicative of a polarization (and/or geometrical reorganization) of the micro-roughness under the effect of bottom currents.

4.6.1.2 Nearshore sandy area - Westdiep swale

Bathymetry and backscatter maps for this area are presented in Figure 4.5B and 4.5E, respectively. The backscatter is relatively homogeneous although a detailed inspection of the ROI indicates slight variations in backscatter values (~ 3 dB) between troughs and crests of the mega ripples. This may be indicative of variations in sediment type (granulometric differences) leading to finer fractions in the troughs and coarser ones on the crests and slopes. Figure 4.8 shows the inverse trend between depth and reflectivity profiles within this ROI. The mega ripples are flood-dominated and are oriented perpendicular to the coastline. In terms of substrate and morphology, this study area can be divided into two distinct sub-areas: the northernmost part (within which the ROI is situated), composed of well- to moderately-sorted fine to medium sand and characterized by flood-dominated mega ripples (λ = ~ 20 m, H = ~ 0.8 m – see Fig. 4.6N) and the southern part (moving coastward), where ripples become progressively smaller (λ = ~ 13 m, H = ~ 0.3 m) evolving into a very flat (< 1°) area, mostly composed of well-sorted medium to coarse sand. Whilst some biological content was present in the northernmost grab samples, considerable amounts of benthic biota were present in the remaining samples. Benthic flatfish, bivalves (*Macoma baltica*, Linnaeus 1758) and abundant (> 10 per sample) echinoderms (*Echinocardium cordatum*, Pennant 1777) and brittle stars (Ophiura sp.) were predominant. High bioturbation characterizes this area which may lead to important modifications of the water-sediment interface over short temporal scales.

The 13-h time series for this site is presented in Figure 6 (second column). In contrast to the very stable Kwinte swale study site, the AR time series for the Westdiep (Fig. 4.6B and 4.6H) present very-high variability throughout all angles reaching > 3 dB for the entire angular sector ($BS_{0\text{-}73°}$) and > 2 dB in the oblique sector ($BS_{30\text{-}50°}$; Fig. 4.6H). The trend observed in BS (Fig.4. 6E) partly follows the oscillation of the tidal level with a significant and progressive (starting from T_8, ~

15:00) decrease in mean BS during the ebbing phase of the cycle. During both flood events values remain stable and fluctuate within a ± 1 dB range. While the backscatter dependence due to survey azimuth was counteracted by the mono-directional survey strategy, a strong dependence to morphology is observable in this study area (Fig. 4.8) and is confirmed by 3D visualization of the mega ripples (Fig. 4.7). A pattern of ripple-cap inversion between flood and ebb tide flows is observed (Fig. 4.6N), leading to build-up of finer material on the stoss side of the ripples (note the change in profile orientation and shape between left and right panels in Fig. 4.7). This is visible in Figure4. 6N where the ebb-phase profile shows an accretion (denoted by the white space between the vertical accuracy envelopes) of ~ 6 cm.

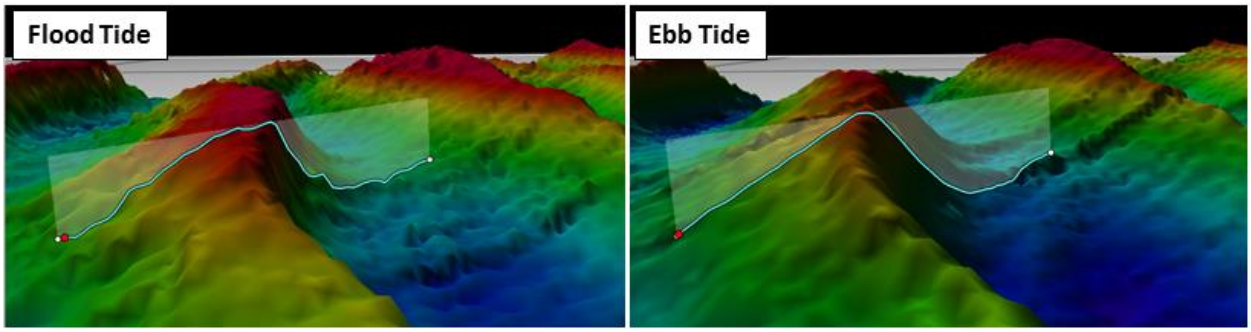


Figure 4.7 - 3D models of a mega ripple found within the ROI of the Westdiep experiment (central ripple in Fig. 6N; same peak flood and ebb times as in Fig. 6K). Vertical exaggeration = 6x. To verify the consistency of this pattern over the entire study area, profiles were extracted from the full transect; different sub-areas of the entire transect and at different angles i.e., nadir, oblique and fall-off angular regions of the swathe (not shown).

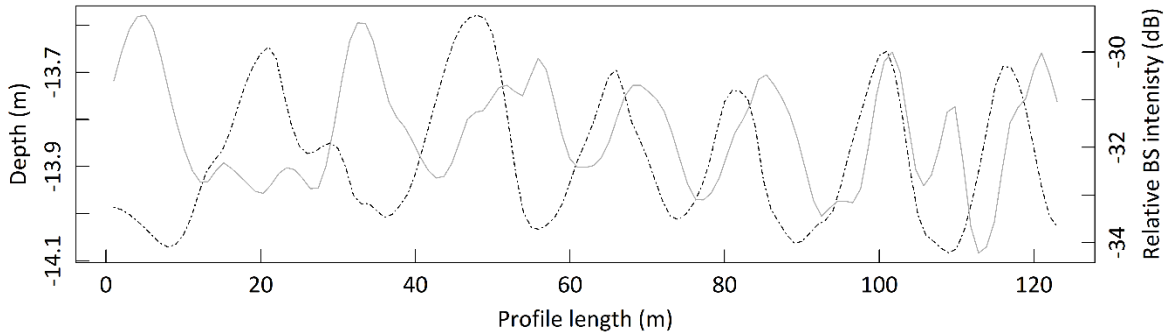


Figure 4.8 - 2D profiles of bathymetry and backscatter extracted from 1-m horizontal resolution raster data within the ROI (Experiment II, Westdiep swale). The grey solid line indicates the depth whereas the dotted black line the backscatter for the same 2D profile. Note the reverse trend in the two profiles. A ~ 3 dB difference between crests (lower BS ~ -33 dB) and throughs (higher BS ~ -30 dB) suggests the presence of different grain sizes along the ripple morphology.

For this experiment, several physical processes were captured by the oceanographic sensors mounted on the benthic lander (Fig. 4.9). They provide ground-truth information to understand the dynamics during the experiment and possibly to explain the observed patterns in the MBES-BS data. Non-parametric correlation coefficients obtained by the Spearman ρ rank method are presented in Table 4.5. While correlation may not directly imply causation, it might be indicative of the processes that drive the variability of the MBES BS at the study site in association to

the hydrodynamic forcing. First, significant correlations between the mean MBES-BS, tidal level (ρ = -0.56, $p < 0.05$) and the current speed (ρ = 0.59, $p = < 0.01$) were found, suggesting that hydrodynamic-related processes played a role in the MBES-BS signal fluctuation. Significant correlations with SPMc at ~2.4 mab (from OBS and LISST sensors; ρ = -0.66, $p = < 0.01$ and ρ = 0.84, $p = < 0.0001$ respectively) were also detected. SPMc was however insufficient to explain the presence of a significant (i.e. > 1 dB) absorption event and these correlations are likely indicative of a similarly fluctuating behaviour of the variables. Continuing, the vertical current velocity (in the z axis measured at 0.2 mab) and the alongshore current vector were also significantly correlated to the mean MBES BS with ρ = 0.75, $p = < 0.001$ and ρ = 0.58, $p = < 0.01$, respectively. This could be explained by the influence of the alongshore hydrodynamic forcing (the cross-shore correlation was weak and not significant) on the sand transport at the boundary layer, modifying the geometry of the bedforms and thus the resulting mean backscatter. Seabed altimetry (measured by the ADV sensor at 0.2 mab) correlated with ρ = 0.54, $p = < 0.05$.

Table 4.5 - Correlation matrix obtained by the Spearman rank method (lower triangle shown). Significance levels of the correlations are denoted by asterisks: Legend of the significance in the bottom row of the table. Values in italic = > 0.7.

Variable/Spearman	Mean MBES BS
Tide level	-0.56*
Curr. speed	0.59**
ABS D50 (1 mab)	0.24
ABS SPM (1 mab)	-0.38
OBS SPM (2.4 mab)	-0.66**
LISST Trans. (2.4 mab)	0.84****
ADV curr. (Z)	0.75***
ADV curr. cross-shore	-0.2
ADV curr. alongshore	0.58**
ADV altimetry	0.54*

* Significance: p < .0001 '****'; p < .001 '***', p < .01 '**', p < .05 '*'

The tide-level trend over the duration of the experiment is reported in Figure 4.9A, along with its corresponding current velocity. In this area, the amplitude of the spring tidal range is around 5.42 m with both ebb- and flood-peak tidal phases having velocities greater than 0.4 m/s, which can resuspend material (Soulsby, 1997). Van Lancker (1999) estimated the median particle size able to be resuspended and transported by subtidal alongshore flood and ebb currents in this area being respectively 420 µm (medium sand) and 177 µm (fine sand) under the spring tidal regime. The NE-directed alongshore current vector (Fig. 4.9B) is the dominant component of the flow in this study and is the main driver of sediment mobility and geometrical reorganization of the micro-roughness. This is illustrated by the tidal ellipse (Fig. 4.9C) which presents a SW-NE elongated shape. The vertically-averaged ABS SPMc (for the 1 mab profile; Fig. 4.9D) is in close agreement with the tidal level where highest concentrations are observable during both flood tide events

(Fig. 4.9A) reaching peak current velocities of up to 0.6 m/s in the alongshore direction (Fig. 4.9B). Potential of deposition/erosion events during the experiment may be assessed by the combined observation of the D50 vectors (from LISST and ABS - Fig. 4.9F and 4.9G respectively), seabed altimetry (Fig. 4.9H), and the alongshore current (Fig. 4.9B). During the first slack water window (around 16:00), larger median grain sizes in the suspended sediment are detected reaching ~160 µm and 220 µm respectively for ABS and LISST sensors (Fig. 4.9G and 4.9F). In the following ebb phase (~19:00), under a significantly weaker alongshore ebb current velocity of about 0.2 m/s, the suspended finer matter may aggregate, sink and settle to the bottom, remaining trapped until the next flood phase (particularly considering the flood-dominated orientation of the study area and the steep lee side of the mega ripples), leading to a ~ 2 cm difference in seabed altimetry (Fig. 4.9H) and a slight increase in turbidity during the ebb tide (note the OBS SPMc peak around 19:00 in Fig. 4.8E). While this study site is situated beyond the far-field of the turbidity maximum zone, pre- and in-survey meteorological conditions induced a rather turbid ebb flow compared to the flood-incoming water masses. This may possibly introduce fine matter residue into the sandy system (Fettweis and Baeye, 2015). Nevertheless, Figure 4.9I indicates that throughout the experiment, the water column at ~ 3 mab (and presumably above this level and up to the surface) was very clear with maximal SPMc of ~ 0.05 g/l.

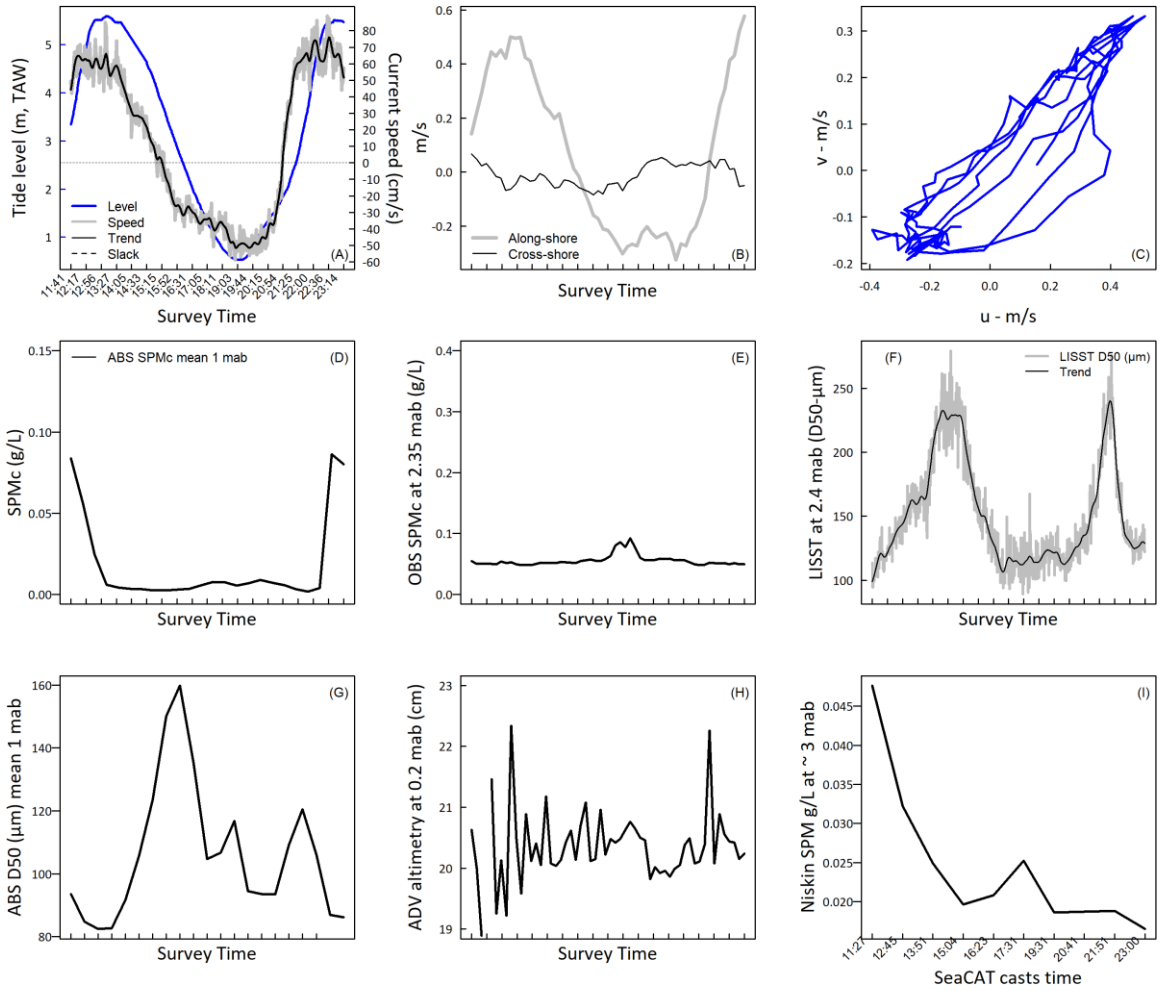


Figure 4.9 - Synthesis of the benthic lander dataset of the Westdiep area (second experiment). A) Tidal level with current speed. Slack water indicated by the horizontal dashed line. The trend of the current speed is achieved by fitting of a cubic smoothing spline function: B) Current speed in along and cross-shore directions; C) Tidal ellipse for the duration of the experiment; D) Vertically-averaged SPMc for the 1 mab, as detected by the ABS sensor; E) Same as (D), but detected by an OBS installed at 2.35 mab; F) Median particle diameter (D50) detected by the LISST at 2.35 m; trend obtained as in (A); G) Vertically-averaged D50 as in (D); H) Seabed altimetry from an ADV sensor at 0.2 mab; I) SPM ~3 mab, obtained from the water filtrations of the CTD-installed Niskin bottle.

4.6.1.3 Nearshore muddy area - Zeebrugge, MOW 1

Bathymetry and backscatter maps for this area are presented in Figure 4.5C and 4.5E, respectively. The substrate type here is muddy sand with the sand part being < 200 µm (fine sand). The bathymetry is very flat with < 30 cm depth difference within the ROI (Fig. 4.6O). Both in the backscatter and bathymetry images there is evidence of bottom trawling, resulting in regularly-spaced striped depressions all over the area. In the immediate proximity of these trawl marks, erosional features appear as relatively small (5 to 15 m in diameter and ~30 cm in depth) concentric/elliptical scours, corresponding to patches of substrate being eroded and washed from the bed likely as a direct consequence of fishing gears' passage enhanced by local hydrodynamic forcing.

The 13-h backscatter time series for this area is presented in Figure 4.6 (third row). Similarly to the Westdiep site (2nd experiment), the average backscatter fluctuates significantly beyond the ± 1 dB sensitivity threshold and a trend consistent to the tidal oscillation is observable (Fig. 4.6F). This study area reaches the highest level of variability: the envelope of variation exceeds 4 dB at 45° and respectively 5 and 7 dB in the specular and fall-off regions (Fig. 4.6C, 4.6I). Higher BS averages occur around the end of the first ebb (~23:00 – 00:00) and around peak time of the second ebb phase (09:00 – interestingly this occurs in concurrence to the higher percentages of sand fraction in the Reineck samples shown in Fig. 4.10A and the strongest ebb current > 0.5 m/s). Lower BS averages occur noticeably during the second ebb tide phase, at around slack water time (~ 08:00).
The interquartile range of the backscatter is about 2 dB (Fig. 4.6F). Comparing angular responses from peak ebb and flood tidal moments (Fig. 4.6L), no azimuthal dependence is detected (no changes in shape) confirming the absence of organized roughness in this flat area (see the 2D profiles in Fig. 4.6O). Despite the shape of the curve remaining unaltered between ebb and flood, differences > 2 dB are observable across the full angular range (i.e. a general decrease in reflectivity; Fig. 4.6L), suggesting the transition of this seafloor patch to different states during different phases of the tidal cycle. A set of ground-truth data is presented in Figure 10 to help interpreting the MBES-BS time series. Figure 4.10A shows the fine sand (<= 200 µm) and mud (<= 63 µm) fractions from the first centimetre of the time series of sliced Reineck box core samples (12 samples, 1 approx. every hour). The tide level (blue line) is superimposed together with the corresponding current velocity (black line - from an ADP sensor). During the two ebb-tide phases, prior to slack water, the sand fraction in the samples is globally more important than during the flood tide where, in concurrence to a decrease in current velocity, samples are dominated by mud (up to ~75% content).

Figure 10 B shows the bi-temporal image differencing change detection between maps of 21st and 24th November 2017 (pre- and post-experiment) summarized into 3 categories of persistence and from-to transitions between mud and sand fractions. While persistence is the dominant component of the change, the sand-to-mud change is observable at the central part of the study area where it forms an elongated pattern (where the bathymetry presents a slight channelling depression compared to the surrounding). The mud-to-sand pattern appears as more randomly distributed, forming patch-like features.

In Figure 4.10C, SPMc from the OBS chain, and the mean MBES BS and QF of the ROI are displayed. Again, the MBES BS acquires a most absorbing character when the SPMc reaches its maximum (around 08:00; ~2.8 g/l at 0.3 mab, 1.3 g/l at 1 mab and ~ 1 g/l at 2.4 mab) and reversely. The relationship between the mean MBES BS and the near-bed SPMc can be captured by a least-square linear regression (R^2 =0.47, $p < 0.01$) that is significant, as well as by the Spearman correlation coefficients (Table 6). Visualization of these data (Fig. 4.10C) indicates that the least

accurate sonar bottom detections (red line) occurred in concurrence with the highest SPMc (particularly at 0.3 mab), resulting in the lowest BS averages. Oppositely, during the flood phase of the tide (~ T13 to T18 – 01:00 – 03:00) the accuracy of the bottom detection increases with decreasing SPMc. This suggests the presence of a dynamic high-concentrated mud suspension (HCMS) which, once settled, increases the volume of the water-sediment interface (forming a “fluffy” layer which increases the burial volume of the seafloor surface) to which the registration of bottom detection and echo intensity are sensitive to. As such, under this configuration, the active seafloor target considered in bottom detection will change from an extended surface (i.e., the relatively “clean” seafloor surface), to a volume cell (i.e., a “slice” or a truncated prism) populated by point-scatterers, which may raise or attenuate the BS level (Lurton, 2010). The behaviour of this HCMS layer appears as the dominant driver of variability of the MBES-BS time series of this area, leading to short-term and progressive changes in scattering mechanisms (i.e. from a relatively “clean” surface with > 50 % of sand to a relatively “fuzzy” mixed sediment interface topped by a ~30 cm deposition of fluffy material).

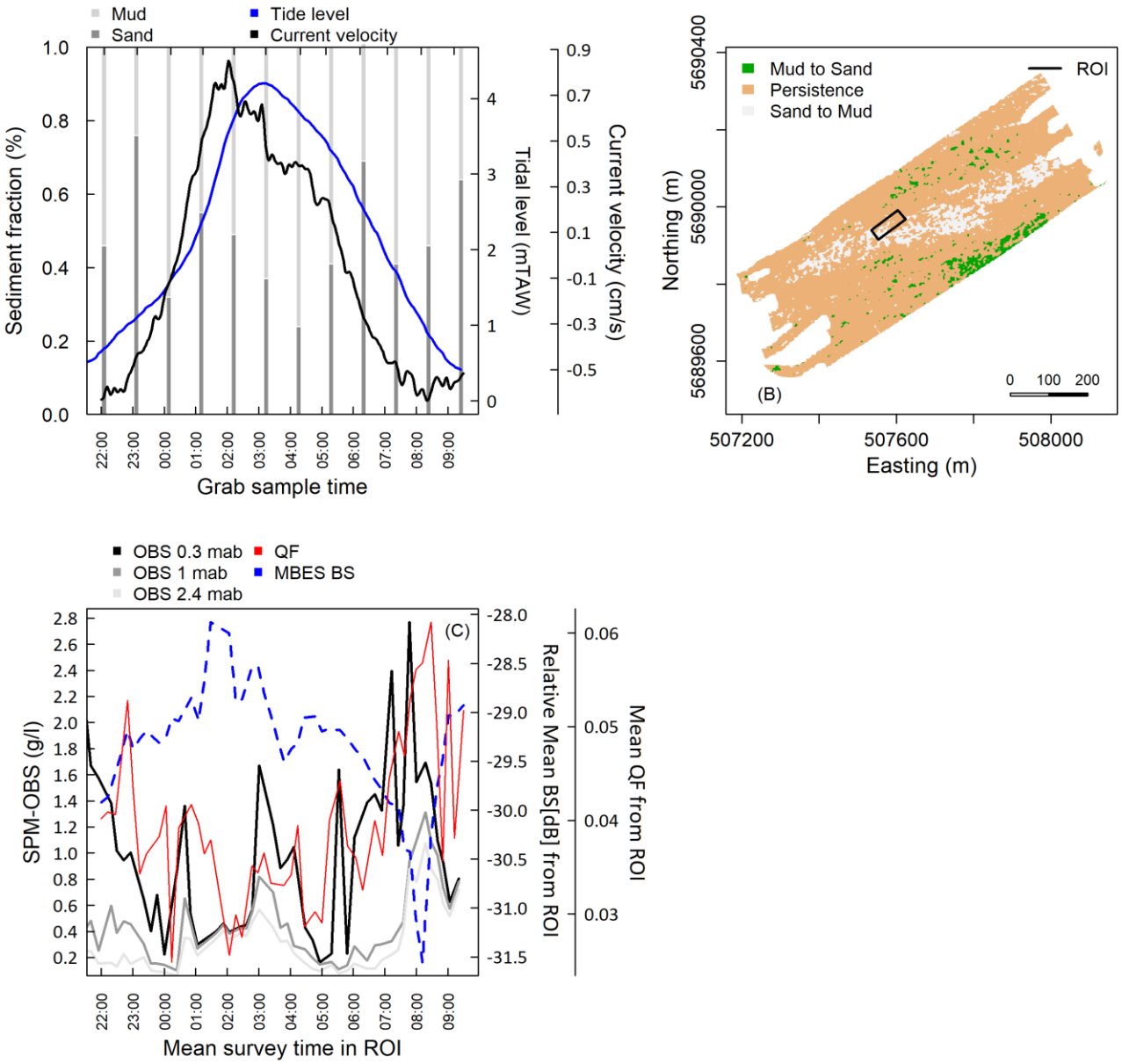


Figure 4.10 - A) Variation in particle size of the first centimetre of the Reineck box-cores time series (n = 12, collected approximately every hour – the above x axis indicates their real position in respect to the tidal cycle), together with the tidal level and the current velocity (respectively blue and black lines, right axis); B) Bi-temporal image differencing (algebraic) change detection between maps of 21st and 24th November 2017 (pre- and post-experiment) summarized into 3 categories of persistence and from-to transitions. Green: Mud to Sand transition; Orange: Persistence; and Grey: Sand to mud. Black rectangle: the ROI; C) SPMc derived from the OBS sensors chain (continuous lines, left axis), mean MBES BS from the ROI (dashed blue line, right axis) and mean Kongsberg QF (continuous red line).

Table 4.6 - Correlation matrix obtained by the Spearman rank method (lower triangle shown). Significance levels of the correlations are denoted by asterisks: Legend of the significance in the bottom row of the table.

Variable/Spearman rho	Mean MBES BS
Mean Kongsberg QF	-0.61****
OBS SPMc 0.3 mab	-0.69****
OBS SPMc 1 mab	-0.40**
OBS SPMc 2.4 mab	-0.35*

* Significance: p < .0001 '****'; p < .001 '***', p < .01 '**', p < .05 '*'

4.7 Transmission losses

In this section, transmission losses during the experiments are evaluated. The variability of the seawater absorption coefficient (Francois and Garrison, 1982a, b) was computed based on surface temperature and salinity from SBE 21 SeaCAT Thermosalinograph values stored in ODAS (On Board Data Acquisition System; R/V *Belgica*) and from SBE 21 SeaCAT Thermosalinograph and SBE 38 Sea-Bird Digital Oceanographic Thermometer values stored in MIDAS (Marine Information and Data Acquisition System; R/V *Simon Stevin*) systems. The echo level uncertainty (in dB) was estimated for the average depths of the study sites and for different slant ranges corresponding to nadir (0°), oblique (45°) and grazing (70°) angles (see Table 4.7). The uncertainty magnitudes resulted as negligible (N) for beams at nadir and small to negligible (S-N) for beams at 45° and 70° (according to the nomenclature proposed in Malik et al., 2018).

Table 4.7 - Table reporting the estimated uncertainty introduced by the seawater absorption coefficient (sensu Francois and Garrison, 1982a, b) for each experiment and for nadir (0°), oblique (45°) and grazing (70°) angles. This uncertainty estimate was accounted for during acquisition.

Experiment	Overall α_W error (dB/km)	Depth (m)	0° (dB)	45° (dB)	70° (dB)	Uncertainty score*
Kwinte swale	2	30	0.11	0.17	0.35	S
Westdiep swale	2	20	0.08	0.11	0.23	N-S
Zeebrugge MOW 1	1	10	0.02	0.028	0.05	N

* N = Negligible (0.01 – 0.1 dB), S = Small (0.1 – 1 dB), M = Moderate (1 – 3 dB), H = High (3 – 6 dB), P = Prohibitive (> 6 dB). Uncertainty score nomenclature after (Malik et al., 2018).

For the second experiment (Westdiep area), a set of CTD down-casts allowed investigating in more detail the absorption variability over the water-column profile. Figure 4.11A-D shows the vertical variability of temperature, salinity, sounds speed and α_w for one CTD down-cast: the variability of these measures is within the

instrumental error of the sensors, indicating the high homogeneity of the water column. Figure 4.11E shows the mean values of the vertically-averaged absorption coefficients for each of the 10 CTD casts, individually displayed in Figure 11F. In this environmental setting, the stability of the vertical profiles justifies the use of surface values to correct for absorption during data acquisition.

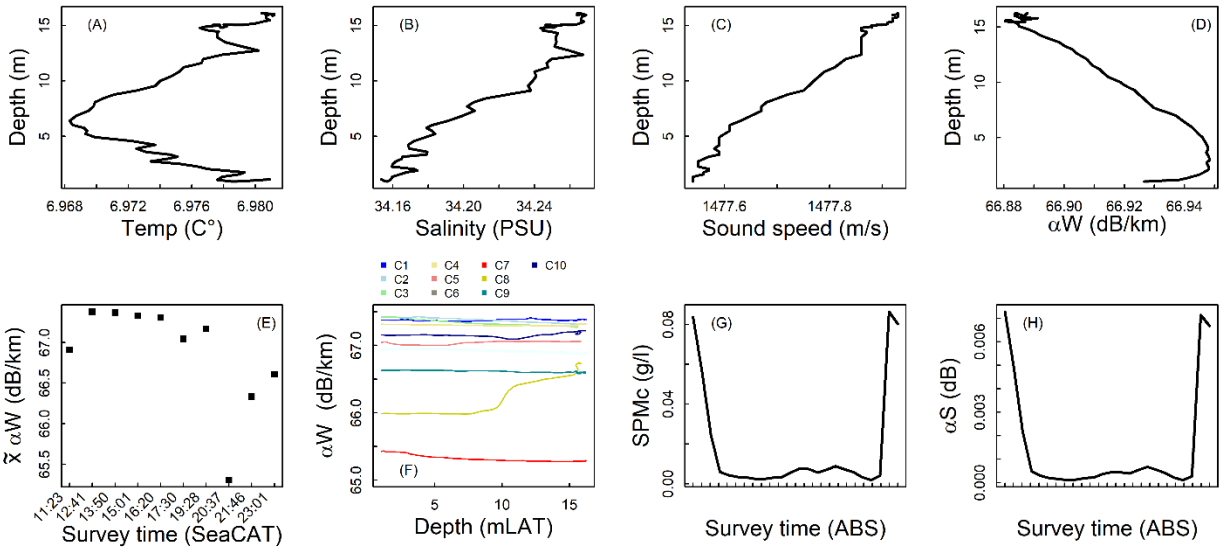


Figure 4.11 - Temperature (A), salinity (B), sound speed (C), and absorption coefficient (at 300 kHz) due to seawater (D) over depth for one CTD downcast (~15 m). Vertically averaged (E) and full profiles (F) of α_w coefficients. G) Averaged SPMc (g/l) for the metre profile above seabed as obtained by the ABS sensor installed on the benthic lander. H) Absorption due to suspended sediment (α_s) for the 1-m profile above seabed computed as a function of vertically-averaged SPMc in G and vertically-averaged grain size (shown in Fig. 7G).

For this second experiment, the SPMc and median grain size (D50) obtained from the ABS sensor (Fig. 4.11G-H) allowed estimation of the transmission losses due to SPMc. Figure 4.11G reports the vertically- averaged SPMc and Figure 4.11H shows the dB loss for the 1-m profile. For this experiment and for such sound travel-path, fully negligible (N) influences of SPMc on the mean BS level are observed.

Nonetheless, to better appraise the uncertainty potentially introduced by this environmental factor, vertical sediment profiles (approximating to the full travel path of the acoustic signal and defined in a conservative way maximizing the SPM impact) were reconstructed for the second and third experiments as specified in the “Materials and Methods” section. As in Table 4.7, Table4. 8 reports the estimated transmission losses for nadir (0°), oblique (45°) and fall-off slant ranges (70°).

Table 4.8 -. Table reporting the estimated uncertainty introduced by the suspended sediment absorption coefficient (α_s sensu Richard et al., 1996; Urick, 1948) for the 2nd and 3^d experiments and for nadir (0°), oblique (45°) and grazing (70°) angles. Out of the four and two profiles for the 2nd and 3rd experiments respectively, only the worse-case scenarios are shown.

Experiment	Depth (m)	0° (dB)	45° (dB)	70° (dB)	D50 Upper/Lower (µm)	Uncertainty score
Westdiep swale	15	0.13	0.18	0.38	100/100	S
MOW 1	10	0.35	0.48	1	63/125	S

Transmission losses due to suspended sediment remain small for both experiments and for the depths, concentrations and particle sizes assessed. Noticeably, for the second experiment in the sandy and clear-water area (Westdiep swale), losses due to seawater only and those due to suspended sediment show similar magnitudes and increasing the D50 in the lower part of the water-column causes little changes. Oppositely, for the third experiment in the maximal turbidity zone, the echo level attenuation increases significantly reaching up to ~ 0.5 and 1 dB at oblique and fall-off slant ranges respectively, showing slight increases with increasing particle size.

4.8 Discussion

Mapping for monitoring requires repeated measurements of the same seafloor areas over short-, medium- to long-term time scales (i.e. diel to decadal time scales). Three field experiments were conducted in the BPNS under spring tide regime to investigate the short-term effect of environmental sources of variance on the acoustic signature of predominantly gravelly (Kwinte swale), sandy (Westdiep swale) and muddy areas (Zeebrugge, MOW 1). These field studies were also aimed at appreciating the sensitivity of the MBES-measured BS to relatively subtle variations in the nature of the water-sediment interfaces at stake. The backscatter time series were analysed, and the signatures and trends were related to seabed physical properties measured in situ, using several approaches. The potential sources of short-term (half-diel) variability that were investigated relate to: roughness polarization and morphological changes, water-column processes (transmission losses due to seawater and suspended sediment), and surficial substrate changes.

4.8.1 Short-term backscatter tidal dependence

The MBES-measured BS variability and its causes differed considerably between the three investigated areas. Overall, the effect of water-column absorption variability (i.e. due to seawater only), was ubiquitously negligible to small; this was expected given the shallow depths surveyed and the good instrumental control of the local seawater characteristics. The effect of suspended sediment on the transmission

losses can be expected to cause little uncertainties in the sandy and gravelly areas outside the turbidity maximum zone in Belgian waters; it could however become moderate to prohibitive in deeper areas or in case of dense plumes of sediments in the water column related to human activities (dredging, trawling). In general, considering jointly the seasonal and spatial variations of SPMc in the BPNS (Fettweis et al., 2017), a maximal water depth of ~ 50 m over the region, and the preliminary observations from this investigation, it may be surmised that for the gravelly and sandy clear-water areas (offshore and in the SW nearshore areas), the effect of suspended sediment will always be small since the highest volume concentrations are to be expected in the lowest layer of the water column, thus involving too short a sound travel path to significantly affect the echo level. Previous investigations on the effect of near-bed SPM on BS for the first study area can be found in Roche et al., (2015) and De Bisschop (2016) and reported negligible effects. On the contrary, in the nearshore zone with soft-material sediments and maximal turbidity, significantly higher volume concentrations can be met even in the upper part of the water column, evidencing the importance of this SPM-caused attenuation even at very-shallow depths (10 m). Besides these environmental factors, the envelopes of variability were mainly driven by short-term successional changes of the underlying morphology and of the water-sediment interface physical status, thereby relating to actual changes in the targeted seafloor.

4.8.1.1 Experiment 1 – Offshore gravel area

Overall, the results pointed at a high stability (< 0.5 dB excluding nadir beams in the angular range 0-18°) of the Kwinte gravel area. This was expected given the known bathymetric and sedimentological spatio-temporal stability of this area (Roche et al., 2018). This good stability is explained by year-round, well-mixed and clear water masses (van Leeuwen et al., 2015) and possibly by an overall stochastic re-organization of the substrate (i.e. geometric micro-changes of the sand and bioclastic material) configuration under the effect of currents which limits significant alterations of the interface backscatter. The backscatter AR was here a particularly useful measurement, not only to gain a physical understanding of the backscattering characteristics of the substrate type (the AR curves show three distinct shapes characteristic of each substrate type; see Fig. 4.6A-C), but also to detect the presence of a weak azimuthal-like dependence thanks to the BS values measured in the steep-angle range (see Lurton et al., 2018). This would have been impossible using solely backscatter mosaics which by nature lack the angular component (as the change detection carried out in Rattray et al., 2013 and Montereale-Gavazzi et al., 2017). This shows that a compensation of mosaicked backscatter imagery using an angular interval in the range 30° - 60° (as in e.g. FMGT standard processing) would omit the azimuthal dependence (which in this gravelly/hillocky terrain extended only up until 18°) while assessing changes of interest (i.e. sediment type at oblique angles) within such seafloor type.

4.8.1.2 Experiment 2 – Nearshore sandy area

The sandy area in the nearshore Westdiep swale showed significant variability (> 2 dB at 45° and > 3 dB over the full angular range) for the time assessed. Water-column processes caused here also negligible impact. Here, most BS variability was best explained by azimuthal dependence, similarly to studies in other sandy/siliciclastic areas (Fezzani and Berger 2018; Lurton et al., 2018). Ripple features are predominant in such areas (Masselink et al., 2007) and, under the effect of both flood and ebb currents, a geometric reorganization of the morphology at various scales may occur. Wave-induced cross-shore currents, creating micro-ripples, may further contribute to MBES-BS variability: when these ripples are perpendicular to the sonar across-track acoustic line of sight, MBES-measured BS may be altered significantly (e.g. Briggs et al., 2001; Richardson et al., 2001). Besides the azimuthal dependence normally limited to steep angles (Lurton et al., 2018), significant variability was also observed at angles beyond 40° (i.e. > 2 dB at 45°), suggesting that some degree of sedimentary changes for the period assessed did occur. As in Ernsten et al., (2006), ground-truth observations were indicative of changes at the interface that likely resulted from cyclicity in deposition/erosion events. The contribution of biological activity (i.e. bioturbation) was not quantified here but is also expected to increase the BS variability. Considerable amounts of biota were observed surrounding this study site which align with previous studies (Van Hoey et al., 2004; Degrear et al., 2008). Feeding and burrowing behaviour of certain benthic species can lead to drastic modifications of the sediment in terms of its geotechnical composition (e.g. permeability, porosity, compactness and roughness; Rowden et al., 1998) and can therefore have large effects on the backscatter level by altering the average water-sediment impedance contrast. Furthermore, presence of individual species per se can act as surface scatterers: e.g. Degraer et al., (2008) and Holler et al., (2017) related part of the high-reflectivity facies in their acoustic maps to the widespread presence of respectively the tubicolous polychaete L. conchilega and the brittle star A. filiformis modifying the micro-roughness as a function of their feeding behaviour (rising of tentacles in the water-column/boundary layer). Recently, laboratory tank-based experiments showed that in sandy sediments the effect of microphytobenthos photosynthetic activity can also introduce a variability of the backscattering properties of the inhabited marine sediment by as much as ~2.5 dB at 250 kHz and over a diel cycle (Gorska et al., 2018). This experiment demonstrates the necessity to jointly analyse mosaicked and AR BS to avoid misinterpretations of the observed changes, particularly in sandy/siliciclastic areas such as on this Westdiep area. It is worth noting that this polarization effect may raise specific (and usually underestimated) challenges when merging surveys acquired in different orientations and it will have to be considered in the compilation of existing backscatter maps.

4.8.1.3 Experiment 3 – Nearshore muddy area

The MBES-BS dataset acquired near the Zeebrugge MOW 1 Pile area was by far the most variable, with a mean variability > 4 dB at 45° and beyond for the remaining angular range. The variability was fully unrelated to the azimuthal dependence since the study area ground-truthing showed a levelled and relatively homogeneous terrain, lacking organized morphology. Our interpretation is rather that the variability related to a combination of the intrinsic dynamic nature of the boundary condition (creating a “fuzzy” boundary layer), to granulometric changes at the water-sediment interface (implying fluctuating fractions of sand and mud) and to a highly-turbid water column. This very dynamic muddy/sandy substrate site is particularly complex from an acoustic perspective since the sediment structure exhibits high vertical heterogeneity (i.e. an intricate layering of intercalated sediment matrix of sand and mud on anoxic mud, topped by depositions of up to 30 cm of fluffy material at specific tidal moments). This likely resulted in volumic contributions (i.e. subsurface sediment scattering) as oppositely to the other two experimental sites where the impedance contrast of the water-sediment interface was significantly higher due to the presence of coarser substrates (i.e. gravel, shells and sandy-quartz sediments) and hence dominated the backscattering process. Significant acoustic penetration into the soft sandy sediments is expected to be about 2-3 cm at 300 kHz (Huff et al., 2008), increasing with softer, muddy and unconsolidated sediment as shown in this third experiment. The vertical complexity of the upper sediment layer in this area changes under the influence of the local hydrodynamic forcing that may modify at least the first 3 cm of the interface (observed from analysis of the Reineck box-core data; not shown here), as well as being subject to HCMS dynamics, that can add up to ~30 cm of fluffy material to the seafloor interface (Baeye and Fettweis, 2012). As such, different water-sediment interface configurations progressively occur during different phases of the tide and thus the echo contributions coming either from the upper layer (interface) or from the buried interface (subsurface) will together affect the bottom detection, yielding to shifts in the AR and mosaic values retrieved during the various instances. The accuracy of the bottom detection upon which depth registration relies obviously depends on how “clean” an interface is. To test this observation, the mean Kongsberg Quality Factor (QF) was processed within the ROI to complement the interpretation of the MBES-BS trend. Significant and interrelated associations were found between registered MBES BS, QF and SPMc at 0.3 mab, confirming the MBES-BS sensitivity to the boundary dynamics of this study site, as identified in Fettwes and Baeye (2015) and Baeye and Fettweis (2012).

4.9 Recommendations on future experiments on MBES-BS variability

When tidal dynamicity and/or environment seasonality are expected to cause seafloor BS variability, field studies are recommended to evaluate the significance. While the instrumentation and set-up used in the present investigation proved highly valuable for targeting this aim, some improvements could be brought. Hereafter,

good practice is reiterated, and shortcomings flagged. Future solutions could come in from new instrumentation and/or methodological approaches. In any case, it is critical that the surveys are conducted under favourable hydro-meteorological conditions. Table 4.9 shows the motion sensor-related variables during the data logging. These are used as a form of quality control on the datasets.

Table 4.9 - Motion sensor derived variables used as a form of quality control on the datasets. The sea state (using the World Meteorological Organization scale) is also reported Figures in degrees (°).

	Mean Roll + Range	**Mean Pitch + Range**	**Mean Heave + Range**	**Mean Heading + Range**	**Sea State***
Exp. 1	0.6-0.5	0.3-0.24	0.006-0.03	204-12	2–0.1 to 0.5 m (Smooth wavelets)
Exp. 2	0.8-1.16	2.9-0.12	0.3-0.06	67.4-4	2 to 3–0.1 to 1.25 m (Slight)
Exp.3	3.2-0.65	1.2-0.22	0.007-0.08	60.6-12.3	1 to 2–0 to 0.5 m (Calm to Smooth wavelets)

* World Meteorological Organization code and information of the wave height and appearance.

Noticeably, an average difference of 12° in the heading range during the first experiment could explain the slight azimuthal-like dependence observed. A similar heading average range in the third experiment has no effect in terms of azimuth given the very flat (level) seafloor. At all times during the surveys, the wave height was always lower than 1 m.

Overall, the mono-directional survey strategy applied here was optimal in preventing (or at least minimizing) the effect of survey azimuth relative to the navigation heading (Lurton et al., 2018). Deviations from the planned track-line did occur for a range of reasons but were kept minimal during the experiments. Experience showed that shorter track lines (about 1-km long) were needed to get high-density datasets enhancing the comparability and detectability of trends in the MBES BS and environmental data (n = 44 instances during 3rd experiment, compared to n = 15 and 19 for the 1st and 2nd experiments, respectively). The use of a benthic lander device proved promising, combining various sensors on a single frame, thus retrieving multiple and relevant oceanographic data at once. However, several limitations were identified. First, there was a difference in retrieval location between the oceanographic data and the MBES data. The time bias between the measurements could in part be overcome by coupling the various data types by a unified mean time stamp (mean surveying time within the ROIs). However, the validity of this approach depends on the data acquisition periodicity by each sensor, which dictates the representativeness of the averages produced for certain tidal moments. For example, the ABS sensor used in the second experiment recorded data in bursts of 30 minutes, thereby possibly insufficiently capturing the sand transport behaviour at shorter time scale and possibly missing key moments of the tidal cycle (e.g., peak current velocities). Increasing and homogenising (across sensors) the frequency of registration would improve this limitation. For most optimal experiments, it is recommended to anchor the vessel on four points (i.e. port, starboard, bow and stern). This would allow collection of the various data types closer in space, as well as increase the frequency of seabed and water-column sampling by grabs and down-cast frames, thereby improving cost-time effectiveness of the experiment and data inter-comparability. For the calculation of sound absorption due to suspended sediment in the water column, the modelling reported in Richard et al., (1996) was simplified due to the limited data availability and strong assumptions were made when vertically-averaging or homogenizing the profiles. For more adequate modelling and correction of this phenomenon (e.g., in deeper clear waters where small turbidity changes may already significantly alter the BS), future experiments should collect more detailed information on concentration, particle size and vertical distribution (i.e., Rouse distribution (e.g. Zyserman and Fredsoe, 1994). Uncertainty estimates due to this factor obtained by the reconstruction of full vertical sediment profiles showed that, within nearshore areas, transmission losses can vary significantly and have noticeable impact on the interpretation of multi-pass acoustic surveys. An interesting point to consider here is the rapid evolution of capabilities in water-column backscatter (WCB) data collection by modern MBES systems. Similarly to acoustic Doppler current profilers data, WCB can be calibrated against

water samples to create spatially-explicit profiles of SPMc and particle size, providing detailed information from near the sea-surface, down to the sonar bottom detection (Simmons et al., 2010). This raises the possibility to use more representative data and robustly implement sound-loss corrections in dynamic and deeper survey areas. Additionally, this would also be cost-effective and more complete compared to the deployment of benthic landers and associated instruments which are time consuming, labour intensive and ultimately impeding the retrieval of data for the full water column.

The sonar measurements were interpreted in complement to an array of oceanographic measurements (where applicable) relating to local seafloor and water-column processes. They could be quantified by means of different equipment. Besides deploying multi-sensor benthic landers, downcast of the CTD frame allowed characterizing the water-column profile in detail, thus deriving better estimates of absorption coefficients than solely using sea-surface data. Substrate sampling gears such as the Reineck box-core, retrieving relatively undisturbed samples, proved useful to quantify short-term changes of the substrate composition, and core slicing allowed appreciating the fine-scale layering; such an instrument should be used more systematically in muddy/soft sediment areas to fully evaluate the relations between acoustic response and sub-bottom complexity. Regarding the collection of SPMc measurements, different instruments were used. Chains of OBS mounted on a benthic lander proved very useful to understand the differences between SPMc at the boundary layer (i.e. 0.3 mab) and the upper-water column (i.e. ~2.5 mab). However, they do not provide estimates of the particle size, for which a LISST and/or ABS system should be used. In any case, it is recommended that further studies are dedicated to understanding the differences between optically- and acoustically-derived estimates and that their sensitivity to varying particle sizes and concentrations are addressed (as in Hawley, 2004) so that adequacy of the instruments to different environments can be better understood.

4.10 Implications for repeated backscatter mapping using MBES

The short-term backscatter variability is only one aspect to consider when using MBES for repeated BS mapping. For the ultimate goal of merging datasets in space and time from different systems and vessels (e.g. cross-border datasets), careful consideration of multibeam system accuracy and stability, conditioning data repeatability and scaling, are required (Anderson et al., 2008). This starts with standardizing operational procedures, in terms of acquisition and processing, ideally inspired from community-driven experiences (e.g., the GEOHAB backscatter working group – Lamarche and Lurton, 2018). Accuracy of a MBES system is largely dependent on the calibration process, requiring manufacturer-based operations (i.e. providing users with metrology results) and/or dedicated facilities and instrumentation to carry out in-situ calibration (otherwise unfeasible for hull-mounted systems - see Eleftherakis et al., (2018) and Weber et al., (2018) for detailed

considerations regarding calibration). Data repeatability refers to controlling the spatio-temporal consistency of the acoustic data in terms of instrumental and environment-caused drifts. Beyond direct metrological checks using dedicated equipment, instrumental drift can be controlled by repeated surveys over naturally stable areas (e.g. Roche et al., 2018; Weber et al., 2018) and/or fixed platforms, and regular dry-dock maintenance operations verifying the sensor status (as it is the case for the sonar systems used in this investigation). The focus of this paper was rather on the environmental drift that refers here to evaluating the variability introduced by factors that do not directly relate to seafloor substrate type, but to water-column or near-bed sediment transport processes, as well as to target-geometry insonification related issues (i.e., azimuthal dependence, micro-scale roughness polarization). Such knowledge is important both for “snapshot in time” and for repeated mapping applications since improving the links between environmental variables and acoustic responses can improve the modelling and replication of field observations in space and time and enhance the interpretability of acoustic measurements.

It is important to understand the consequences of short-term environmental variability upon the interpretation of longer-term MBES-BS time series. This requires dedicated and specifically-designed field experiments (e.g. this study, or the SAX experiments in Williams et al., (2009), and those advocated in Lucieer et al., (2018).
As shown in this study, tidal periodicity and seasonality call for careful consideration, especially in shallow areas with soft-material seabed and high sedimentary dynamics. Indeed, successive surveys of a same area may provide different information at various time scales (from day to year). In this regard, it is important that the tidal dependence is analysed per MBES-BS time series. Spotting outliers (i.e. abrupt changes in sediment response) will be relatively straightforward in the clear water and stationary areas (such as the Kwinte swale in the first experiment where tidal dependencies were low) since the magnitude of the short-term variance remains within the envelope of sensor sensitivity (i.e. the manufacturer-set ± 1 dB for EM3002 and EM2040 Kongsberg systems – Hammerstad, 2000). On the contrary, the intrinsic “noisiness” (i.e. periodical variability) of the nearshore areas results in a potentially masking/blurring effect, introducing uncertainties due to the status of the water column (i.e. turbidity) or to the “mobility” of the water-sediment interface. Within such areas the stability threshold must be defined contextually in accordance to the underlying sedimentary environment, and a transition in seafloor status can only be detected from a trend analysis on a sufficient number of serial surveys. Direction and consistency of the trend, regardless of the noise envelope, can be a valuable proxy of change and bypass conflicting results from surveys acquired at different tidal and/or seasonal moments. Interpretation of serial backscatter surveys in such environments should largely benefit both from time series of driving variables (collected via the deployment of benthic landers as in Baeye and Fettweis, 2012; Fettweis et al., 2017) and from regional predictive oceanographic models providing local conditions usable for designing monitoring surveys accordingly.

In the third experiment, the observations showed that regardless of the variability or 'noise' on the AR curves (except for that exerted by the roughness polarization in the second experiment), the main shape of the angular behaviour, indicative of a sediment type, remained the same. While part of this variability was related to transmission losses due to suspended sediment, the observed shifts in backscatter values (a decrease in reflectivity across the full angular range between flood and ebb tide moments) was related to HCMS dynamics which changed the water-sediment interface, evidencing the BS sensitivity to short-term and relatively subtle granulometric and volume heterogeneity changes known to occur in this area (Baeye and Fettweis, 2012). The sensitivity of the angular response to such differences in sediment composition (within the same main sediment class) has been observed in several investigations. For example, in Fezzani and Berger (2018), the high sensitivity of AR is particularly clear: AR curves are used as the basis of classification of a large MBES dataset, resulting in an evident within-cluster variability of up to 10 dB at 45°. Further insights can be found in data presented in (Daniell et al., 2015), in which different AR curves are related to varying degrees of percentage cover of coarse clastic material (i.e. shell and gravel scatterers). This suggests that from well-controlled backscatter measurements with sufficient ground truth data allowing detailed interpretation, the derived BS AR curves can capture instantaneous and temporal physical changes in substrate composition. Critical is then to decipher whether the change was naturally- or anthropogenically-driven, requiring knowledge of the magnitudes of short-term and seasonal variability. A priori knowledge on the magnitude of natural variability would largely assist such interpretations.

4.11 Conclusions and future directions

This research focused on the reliability/utility of BS field measurements by ship-borne MBES for the monitoring of the seafloor interface. More specifically, the aim was to study short-term BS fluctuations specifically associated with tidally-induced half-diel variations of the environment. Three experiments were conducted during which BS was acquired together with environmental variables. Results showed that the latter are important factors in explaining variations in the shape and values of the BS-AR curves and the associated imagery, with various impact levels depending on the local sedimentary configuration. Consequently, it is recommended that, beyond further investigations of the different sources of MBES data variability, detailed environmental variables are systematically collected together with settings of MBES and associated devices, as well as application of best practice in survey designing. For users and surveyors operating within tidally-dominated environments (both for mapping and monitoring purposes), such experiments raise a number of points of interest. Assuming a stable sonar system with no instrumental drift and a rigorously standardised acquisition and processing routine, the following observations were made: (1) in relatively stable and gravelly offshore areas, characterized by clear seawater, the variability due to external sources is limited and the BS measurement

confidently relates to the water-sediment interface. (2) In nearshore sandy areas, roughness polarization may occur at various scales (depending on the hydrodynamic forcing) and a joint investigation of BS mosaics and BS-AR data products are needed to confidently discern between these geometrical effects and actual sediment changes. (3) In nearshore muddy and turbid areas, the influence of suspended sediment is prone to be significant and needs to be corrected for, requiring careful sampling and quantitative estimation of water-column processes. In the absence of sampling, interpretation of MBES BS requires a minima knowledge on the variability of environmental processes, from available time-series data and /or high-resolution sediment transport and current models.

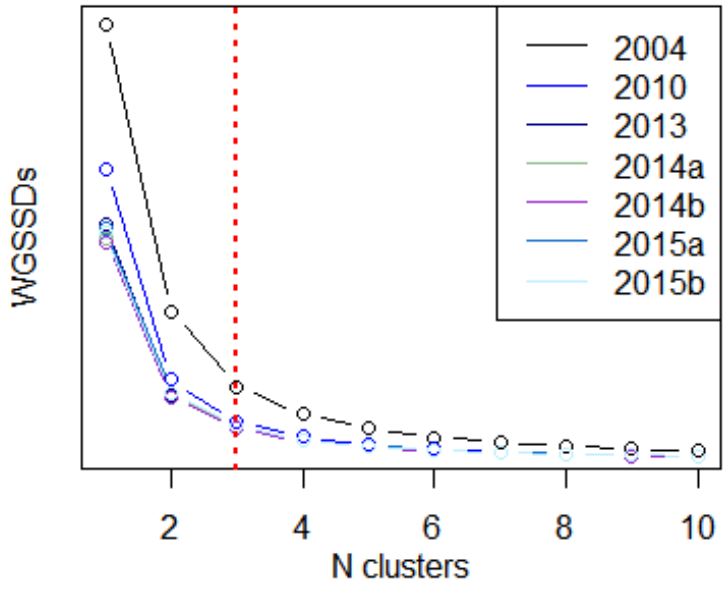
Acknowledgments: We thank the crews, commander and captain of RV *Belgica* and RV *Simon Stevin* respectively for their help during the data collection phase of this investigation. Dr Jean-Marie Augustin of the Underwater Acoustics Laboratory of IFREMER is acknowledged for the help in using the SonarScope MBES processing software.

Author contributions: Conceptualisation, G.M.-G., M.R., X.L. and V.V.L.; Formal analysis, G.M.-G., M.R., M.B., N.T. and F.F.; Investigation, G.M.-G., M.R., K.D., X.L., V.V.L., N.T. and F.F.; Methodology, G.M.-G., M.R., X.L., N.T. and V.V.L.; Project coordination, V.V.L., M.R. and X.L.; Writing—Original Draft, G.M.-G.; Writing—Review and Editing, G.M.-G., N.T., M.R., X.L. and V.V.L.

Addendum – *Errata corrige*

The caption and the interpretation of Figure 4.8 at page 196 have been corrected since the interpretation of the backscatter values in respect to the ripple's morphological features (trough and crest) was erroneous (i.e. inverted).

Chapter 5



5. Seafloor change detection using multibeam echosounder backscatter:

Case study on the Belgian Part of the North Sea

Published as:

Montereale-Gavazzi, G., Roche, M., Lurton, X., Degrendele, K., Terseleer, N., Van Lancker, V., 2018. Seafloor change detection using multibeam echosounder backscatter: case study on the Belgian Part of the North Sea. Marine Geophysical Research 39, 229–247. https://doi.org/10.1007/s11001-017-9323-6

Presented and discussed at:

Conference: **Montereale-Gavazzi, G.,** Lurton, X., Van Lancker, V., Roche, M., 2016. Monitoring seafloor integrity and GES using Hydro-acoustic methods. Estuarine Coastal and Shelf Science Association Conference, Bremen (Germany), 28 November – 2 December 2016 (abstract + presentation)

Invited Workshop: **Montereale-Gavazzi, G.,** Roche, M., Degrendele, K., Lurton, X., Van Lancker, V., 2016. Comparison of Supervised and Unsupervised Benthic Habitat Classifications: Case Study from the Belgian Shelf, ICES meeting - Working Group of Marine Habitat Mapping, Winchester (UK), 11 to 13/05/2016 (presentation)

Conference: **Montereale-Gavazzi, G.,** Roche, M., Degrendele, K., Lurton, X., Van Lancker, V., 2016. Comparison of Supervised and Unsupervised Benthic Habitat Classifications: Case Study from the Belgian Shelf. Geological and Biological Marine Habitat Mapping Conference, Winchester (UK), 2 to 8/05/2016 (abstract + presentation + poster)

5.1 Abstract

To characterize seafloor substrate type, seabed mapping and particularly multibeam echosounding are increasingly used. Yet, the utilisation of repetitive MBES-borne backscatter surveys to monitor the environmental status of the seafloor remains limited. Often methodological frameworks are missing and should comprise of a suite of change detection procedures, similarly to those developed in the terrestrial sciences. In this study, pre-, ensemble and post-classification approaches were tested on an eight km^2 study site within a Habitat Directive Area in the Belgian Part of the North Sea. In this area, gravel beds with epifaunal assemblages were observed. Flourishing of the fauna is constrained by overtopping with sand or increased turbidity levels, which could result from anthropogenic activities. Monitoring of the gravel to sand ratio was hence put forward as an indicator of good environmental status. Seven acoustic surveys were undertaken from 2004 to 2015. The methods allowed quantifying temporal trends and patterns of change of the main substrate classes identified in the study area; namely fine to medium homogenous sand, medium sand with bioclastic detritus and medium to coarse sand with gravel. Results indicated that by considering the entire study area and the entire time series, the gravel to sand ratio fluctuated, but was overall stable. Nonetheless, when only the biodiversity hotspots were considered, net losses and a gradual trend, indicative of potential smothering, was captured by ensemble and post-classification approaches respectively. Additionally, a two-dimensional morphological analysis, based on the bathymetric data, suggested a loss of profile complexity from 2004 to 2015. Causal relationships with natural and anthropogenic stressors are yet to be established. The methodologies presented and discussed are repeatable and can be applied to broadscale geographical extents given that broad-scale time series datasets become available.

Keywords Multibeam, Seafloor backscatter, Change detection, Seafloor integrity, Marine Strategy Framework Directive, Reference calibration area

5.2 Introduction

Human pressures to the marine biome have reached unprecedented extents. Today, globally up to 41% of marine habitats are directly impacted by a multitude of anthropogenic stressors (Halpern et al. 2008). Changes in seafloor substrate composition and spatial configuration may occur as a result of such anthropogenic pressure, but also of natural variability driven by varying hydrometeorological conditions (van Denderen et al. 2015). Our ability to monitor the spatio-temporal dynamics of the seafloor and, ultimately, to relate the observed patterns to driving processes is central to our understanding of marine ecosystems and to the tutelage of the ecosystem services we depend on. This is also recognized in the European Marine Strategy Framework Directive (MSFD—EC 2008-56-EC) in which the seafloor is the backbone of several indicators of 'Good Environmental Status'. For this purpose, seabed mapping, and particularly multibeam echosounding are increasingly used. High-frequency multibeam echosounders (MBES) are considered as the state-of-the-art sonar instruments and are employed by commercial, governmental (i.e. hydrographic services), industry (e.g. oil and gas exploration and exploitation), and research institutions. This is due to the MBES ability to co-register high-density echo time, geometrical features and intensity over large seabed swaths, hence providing depth and intensity data (Kenny et al. 2003). While up until now the bathymetry has been the main focus of hydrographic surveys and mapping programs (i.e. following International Hydrographic Organisation standards of acquisition and accuracy of depth measurements; Wells and Monahan 2002), seafloor reflectivity (backscattered intensity from the seafloor) has only recently attracted interest from a scientific perspective due to its ability to map the water-sediment-interface constituency (Lurton and Lamarche 2015). Mapping this interface over vast areas allows extending information from isolated point locations (in-situ measurements such as grab samples and video observations) to the spatial extent of a digital surface. Moreover, if time series of acoustic data are acquired, it allows the application of change detection methods as developed in the terrestrial sciences with satellite data (e.g. Foody 2002; Pontius et al. 2004; Hussain et al. 2013). This raises the possibility to measure how much the attributes of a particular area have changed between two or more periods. Despite the increasing interest of using MBES backscatter, standards of seabed backscatter acquisition and processing are still under development. A set of guidelines and recommendations was developed by the Backscatter Working Group (or BSWG; see http://geohab.org/bswg) mandated by the Geological and Biological Marine Habitat Mapping scientific committee (GEOHAB). Reaching standardisation of MBES data acquisition and processing procedures is challenging due to the number of manufacturers, multibeam models and dedicated processing platforms, each implementing their own processing algorithms and proprietary software features. This paper addresses the application of change detection methods to capture seafloor substrate changes over a period of 10 years based on a time series of seven datasets of MBES depth and backscatter data (2004–2015). It relates to assessing good environmental status of gravel beds in the

Belgian Part of the North Sea (BPNS) for which the Belgian State specified two indicators on seafloor integrity (MSFD descriptor 6) and for which multibeam technology was put forward as monitoring tool (Belgian State 2012):

1. The areal extent and distribution of the European Nature Information System (EUNIS) level 3 Habitats (sandy mud to mud; muddy sand to sand and coarse sediments), as well as of the gravel beds, remain within the margin of uncertainty of the sediment distribution with reference to the Initial Assessment.

2. Specifically related to the gravel beds it is furthermore specified that the ratio of the hard (gravel) substrate surface area to the soft (sand) substrate surface area must not show a negative trend.

The case study is located within a sandbank system in a Habitat Directive Area of the BPNS. While of high ecological relevance, this area is intensively fished, and marine aggregate extraction started in 2012 near its northern limit. In this paper a methodological framework is presented to assess progress of good environmental status based on multibeam backscatter data. Whilst developed at a local scale, the change detection methodology is promising to be applied on a more regional North Sea level.

5.3 Study area

The study site is approximately 8 km^2 and is located in the proximity of the Western Border of the BPNS, more specifically in the Vlaamse Banken Habitat Directive Area (enacted as of 16th October 2012, EC 92/43/EEC; Fig. 5.1, grey-shaded polygon). It is located in the southern part of a complex sandbank-dune system named the Hinder Banks. Depths range from −8 to −30 m lowest astronomical tide (LAT). Fine to medium sands dominate the sandbank portion of this environment where large and very-large dunes (ranging from 4 to >10 m height) are present (sensu Ashley 1990). The western flank of the main sandbank body forms a transitional area between the bank sandy environment and the adjacent gully. In the latter, medium to coarse sand as well as gravel occur. Gravel provides small-scale structural complexity for ecological successional phases to occur (e.g. deposition of current advected larvae; Houziaux et al. 2007). Seabed maps indicate that the system is very poorly enriched by silt (0–1% silt–clay content; Verfaillie et al. 2006). A series of steep barchanoid dunes is present in the transitional area, with considerable amounts of gravel in the troughs (Van Lancker 2017). Diverse assemblages of sessile and vagile epifauna and benthic fish were observed here in pioneering and more recent studies (Houziaux et al. 2011, and references therein). Hereafter, these are called gravel refugia, since in the majority of the gully epifaunal growth on gravel beds is absent because of severe bottom-trawling occurring since the late 1800s. In gravel areas, these are known to routinely remobilise the gravel clasts (Jones 1992).

Since 2012, a new anthropogenic stressor was introduced in the area, related to sand extraction occurring 2.5 km NE of the Habitat Directive Area. Depending on timing, frequency and amount of extraction and hydrodynamic settings, resuspension of sediment plumes could represent a source of smothering leading to loss of surficial complexity and burial of epifaunal colonies (Thrush and Dayton 2002; Van Lancker et al. 2010; Spearman 2015). To assess environmental impacts, a monitoring programme was setup combining multibeam recordings with seabed sampling, visual observations and water column measurements as well as hydrodynamic and sediment transport modelling (Van Lancker et al. 2016). Sediment plumes arising from the marine aggregate extraction activities, and their deposition, were depicted in acoustic imagery (Van Lancker and Baeye 2015), and numerical modelling results showed that their deposits reach the gravel beds in the Habitat Directive Area up to the study site (Van Lancker et al. 2016). The cumulative volume of marine aggregates extracted throughout the duration of the data time series is shown in Fig. 5.2: larger quantities were extracted from 2012 onwards (~800,000 m^3) to reach a maximum of ~2.4 × 10^6 m^3 in 2014.

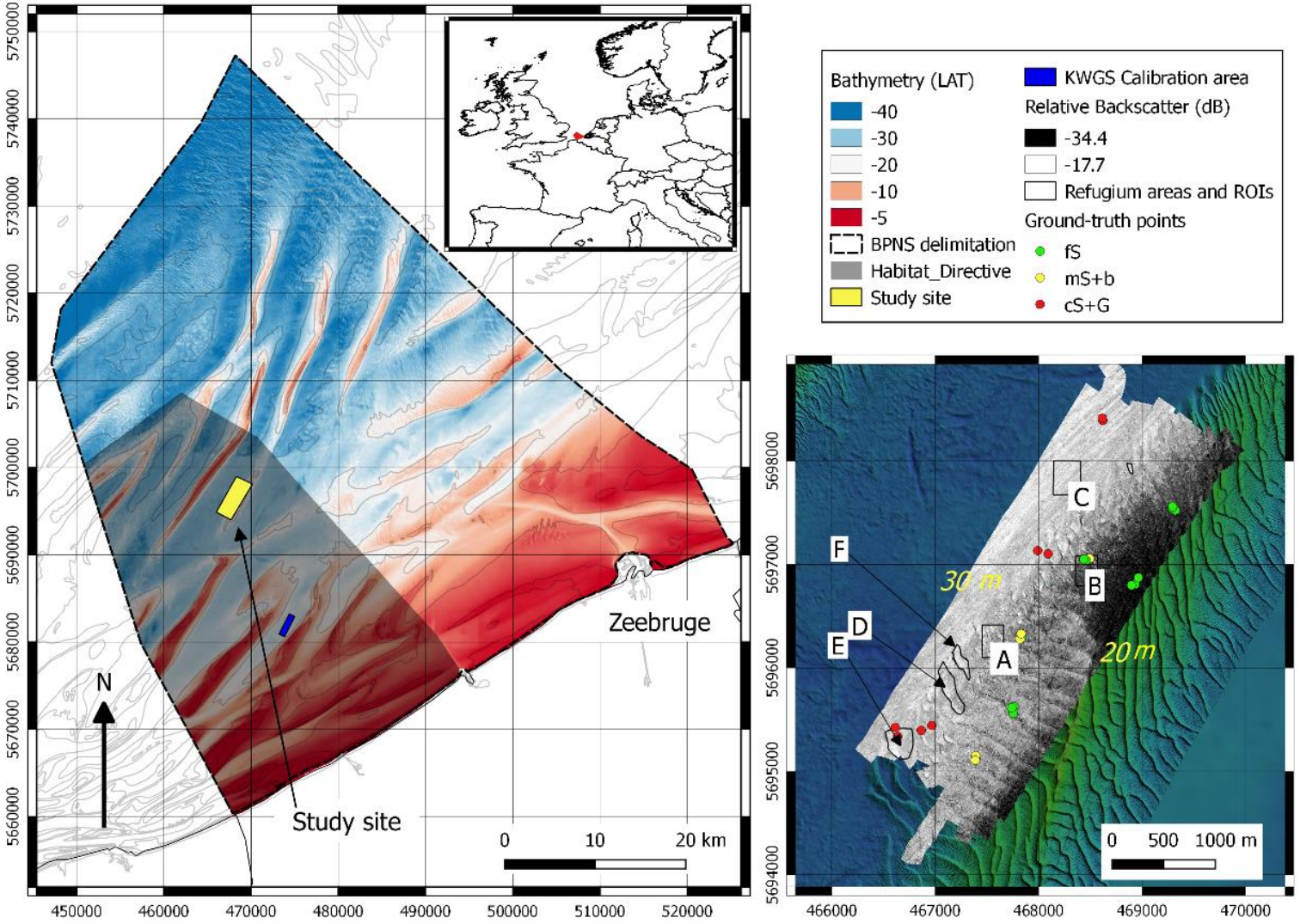


Figure 5.1 - Left Belgian Part of the North Sea (BPNS). Right backscatter (dB) map of the study area with black outline polygons indicating biodiversity rich areas selected as case studies to monitor seafloor integrity

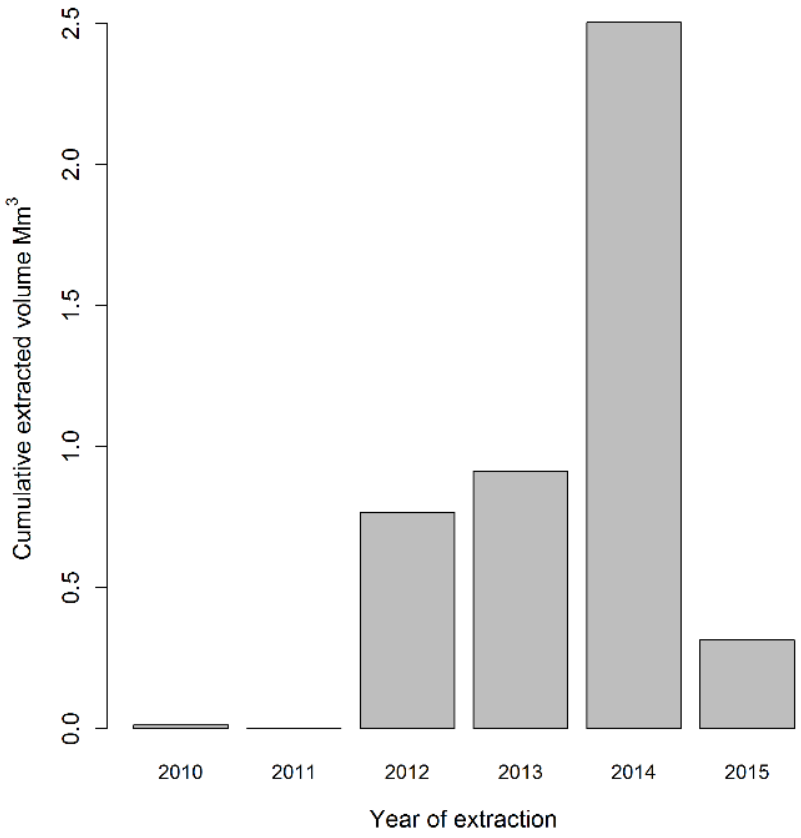


Figure 5.2 - Extracted marine aggregate volume in Mm^3 from Extraction Zone 4 (2.5 km away from the designated area). See Mathys et al. (2011), Van Lancker et al. (2016) for a detailed description on the marine aggregate extraction in this particular area. Effective extraction began in 2012. Data on extraction volumes were provided by the Belgian Federal Public Service Economy, Continental Shelf

5.4 Methods

The “Methods” presents the acoustic and ground-truth data acquisition and processing, comprising a two-dimensional characterization of the spatio-temporal morphological evolution of the seafloor using the bathymetry data and a change detection analysis carried out on the backscatter time series. The steps of the analysis preceding the change detection include the application of supervised and unsupervised classification algorithms and their quantitative comparison. Finally, the change detection using backscatter data is carried out by using both classified (thematic/labelled) and unclassified (relative dB values/unlabelled) backscatter mosaics, as well as applying ensemble approaches.

5.4.1 Data acquisition and processing

5.4.2 Acquisition

The MBES data were acquired by Ghent University in 2004, and later by the Operational Directorate of Natural Environment of the Royal Belgian Institute of Natural Sciences as part of a sand- and gravel-extraction monitoring programme and MSFD-oriented monitoring campaigns (Van Lancker et al. 2016). Of the eight acoustic surveys undertaken between 2004 and 2015, seven were kept for this investigation. Surveys exhibiting a significant amount of navigation artefacts (mostly due to failure in vessel-motion related compensation during rough-sea conditions), were considered unsuitable for the analysis and were discarded. The first survey used a 100-kHz Kongsberg EM1002S, and the remaining six surveys operated a 300-kHz Kongsberg EM3002D (Dual-head system). Both systems were installed on Belgian oceanographic vessel R/V *Belgica*. The hydrographic quality of the EM3002D dataset is consistent with the IHO S44 Special Order, whereas with the former EM1002S only the Order 1A (Wells and Monahan 2002) was attained. Under these standards, the total vertical uncertainty with ±95% confidence levels of the depth measurements result in ±0.63 and 0.33 m vertical error for the EM1002S and EM3002D, respectively, for a depth of 30 m (Tables 5.1, 5.2). These intervals encompass all sources of errors originating from the suite of instrumentation used during acquisition. Pitch, roll, heave and yaw were automatically compensated for during acquisition and a sound velocimeter constantly monitored the sound velocity at the transducers. Survey lines were spaced to reach a good compromise between survey time/costs and quality of the data resulting in a minimum of 20% across-swath overlap between adjacent lines. Throughout the timespan of acquisition (inter- and intra-survey), the MBES settings controlled by the on-board software (i.e. SIS: Kongsberg native acquisition platform) remained unchanged (i.e. pulse length, beam aperture, beam spacing). The state of the antenna transducers was thoroughly checked and maintained for biofouling and deterioration of its components (either by divers or during regular dry-dock operations). Similarly, across all surveys, track lines were sailed in a SW-NE direction. Maintaining operational parameters stable and checking the physical state of the instrument ensured that instrumental drift was kept to the minimum. Regarding sound absorption throughout the water column, the α coefficient (see Francois and Garrison 1982) was computed according to the local seawater properties at the surface which were fed into the acquisition system every half an hour. The necessary water medium environmental parameters were obtained from the Onboard Data Acquisition System (ODAS), which logs these data at 1-s intervals. No vertical profiles of the seawater properties were acquired since in this region the water mass is known to be well mixed throughout the year and no stratification is expected to occur (Luyten et al. 2003; van Leeuwen et al. 2015) and the surface values are considered to be sufficiently representative. To verify instrumental drift on the medium to long term and allow comparison of backscatter

levels in time, data were verified against an area with stable depth and backscatter levels ('KWGS' reference area, blue rectangle in Fig. 5.1). This calibration area (1.8 km2) is located in a gully in-between two sandbanks and is dominated by sand to sandy Gravel. These verifications showed that the oblique incidence backscatter [beam angle sector ±(35°–45°) and ±(0°–70°) for the full angular range] mean values remained, per survey, within 1 dB around the overall mean BS level with no significant trend that would suggest instrumental drift.

Table 5.1 - EM3002D MBES specifications and auxiliary sensors

Parameters	Measure
Central frequency	300 kHz
Number of beams	508 (254/head)
Beam width	1.5° x 1.5°
Beam mode	Equidistant
Angular swath range	200°
Pulse length	150 µs
Positioning systems	GPS Sercell, Furuno and RTK Thales
Motion sensor	Seatex MRU 5
Sound Velocity sensor	Valeport mini SVS

Table 5.2 - Time-series dataset specifications

RV *Belgica* surveys in the Vlaamse Banken Habitat Directive Area			
Survey	Time-layer-ID	Month-year	System
ST2004	T_1	Apr-04	EM1002S
ST2010	T_2	Feb-10	EM3002D
ST1319	T_3	Jul-13	EM3002D
ST1417	T_4	Jun-14	EM3002D
ST1425	T_5	Oct-14	EM3002D
ST1507	T_6	Mar-15	EM3002D
ST1533	T_7	Dec-15	EM3002D

5.4.3 MBES data processing

The backscatter strength (BS) quantifies the amount of acoustic intensity scattered back to the sonar receiver following a complex interaction of the transmitted signal with the seafloor. It is the result of an intricate combination of several physical factors: the seawater-seafloor impedance contrast, the interface roughness and the sediment volume heterogeneity, the signal incidence angle on the seafloor and the acoustical signal frequency (Lurton 2010). Due to the various scattering properties of different seafloor substrates, backscatter can help determine bottom type (e.g. de Moustier and Alexandrou 1991; Hughes-Clarke et al. 1996; Ferrini and Flood 2006) and possibly to infer some of its physical characteristics. However, backscatter data are inherently noisy, showing strong amplitude fluctuations due to the very nature of the scattering process (Lurton 2010), and the possible presence of additive external noise: a first processing stage is to reduce this random fluctuating character by appropriate filtering techniques. A second category of processing aims to correct geometrical artefacts resulting from the characteristics of instrumentation used in the acquisition (i.e. motion and positioning sensors), the seabed geometrical configuration (dictated by the local topography), the velocity and absorption properties of the water medium within which sound is travelling, and the angle of incidence (Lurton and Lamarche 2015). The observed angular response of seafloor backscatter (describing how the reflectivity impact upon echo intensity varies with the incidence angle) can be categorised into three distinct angle sectors. Each are characterized by a different scattering regime (i.e. the specular or near-nadir, the oblique and the low-grazing angle regime), hence they can be treated as separate entities (i.e. statistical populations) (Lurton 2010). In order to produce a sedimentological meaningful image and avoid the along-track banding effect of the three domains, the resulting angular dependence must be compensated. Consequently, the backscatter strength has to be normalised to a conventional reference angle (ideally in the 30°–60° range, but typically 45° is used). Furthermore, several corrections must be applied to the data, in order to account for the sonar

sensor's responses: source levels and pulse length; acoustic transmission losses due to spreading and absorption; 3-D beam directivity patterns; sensitivity of the receiving arrays and electronics; and real-time time varying gain (TVG) corrections applied by the sounder. These various points were addressed in the real-time data reduction scheme applied in Kongsberg Maritime echosounders and during acquisition (Hammerstad 2000). To allow consistency in the last phases of the data processing (i.e. mosaic production) and hence enable their subsequent inter-comparability (in terms of relative dB values expressing a reflectivity scale according to a common reference), the EM3002D data were subject to a standardised processing procedure following the BSWG recommendations (see Lurton and Lamarche 2015). Fledermaus Geocoder (FMGT, v7.4.5.b) and QPS QIMERA (v1.2.4.429a) software suites were used to process the MBES raw data. Initially, tide-corrected bathymetry was produced and exported as 1-m horizontal resolution raster (32-bit float files) and as sound density files for integration in FMGT. The bathymetric surfaces are used to correctly allocate the backscatter snippet traces from single pings to their true seabed position. Each survey was normalised by applying a flat angle varied gain (AVG) filter with a window size of 300. In order to weight nadir pixels and reduce their banding effect, the "No Nadir if Possible 2" algorithm and "50% line blending" FMGT options were applied. As such, the final dataset consisted of (1) relative (standardised to a common reference surface area) backscatter reflectivity (in dB), and (2) bathymetric surfaces (m) at 1 m horizontal resolution. The EM1002S data did not prove to be comparable in terms of backscatter levels with those from the EM3002D system, due to the differing intrinsic properties of the sensors (i.e. electronics and hardware) and to the absence of a cross-calibration of both sensors. Consequently, the first campaign was not included in the pre- and ensemble classification analyses (in "Pre-classification" and "Ensemble approach classification").

5.4.4 Ground-truth data

The ground-truth data used in this study were acquired in complement to the T7 survey. Collection of ground truth is necessary to validate the assumptions developed during the observation of acoustic data and ultimately to derive confidence metrics expressing the validity of the map produced. Ten samples were collected using a Van Veen grab, each with three replicas to ensure the spatial consistency of the acoustic theme being sampled. Video samples were acquired by means of a dropframe, equipped with underwater lights and a camera with a 1 × 1 m field of view. Video-frame data with poor visibility (i.e. due to turbidity or too strong current) were discarded. Visual sampling was very useful to acquire data in the gravel areas where conventional gears failed (i.e. box core and Van Veen). All sample coordinates were corrected for the DGPS antenna layback accounting for the main source of positional error and were mapped with a 10 m buffer. Sample types were described by combining visual and expert observations with grain-size parameters calculated by a MALVERN Mastersizer 3000 instrument. To validate the

consistency in terms of sediment classification versus backscatter levels, the classes' description was compared to previous substrate classification studies within the same area (Roche 2002). Only features visible at the seafloor were described and classified into three thematic classes summarizing the main substrate composition: (1) homogeneous well-sorted fine to medium sand (fS); (2) moderately sorted medium sand with bioclastic detritus (mS + b); and (3) medium-to-coarse sand with gravel clusters (cS + G; Fig. 5.5.3). As will be shown later (see Fig. 5.5 in "Supervised map of the study area"), the fS and mS + b classes are texturally and sedimentologically similar with an overlap in terms of dB ranges, mS + b being a subset of the fS class. This is likely explained by the presence of bioclastic detritus and a significant roughness in the mS + b class which lead to interface scattering having a significant contribution to the overall acoustic return and causing a relatively high level (≈−27 dB) of mean backscatter. On the contrary, the fS class, which is almost entirely distributed on top of the sandbank (in the most dynamic part of the study area, likely with a higher water content than the flank and gully areas) is very well sorted and homogenous, with little interface roughness and no surface scatterers, resulting in the lowest values (≈−31 dB) of mean backscatter (Fig. 5.5). Conversely, the cS + G class features the highest content of coarse material with sparse individual strong scatterers, and high roughness at the interface; hence it corresponds to the highest values (≈−22 dB) of mean backscatter (Fig. 5.5). The described samples were separated into training (2/3) and validation (1/3) sets (Table 5.3). Sample representativeness was assessed visually by plotting the backscatter cumulative frequency distribution of the study area and for the mean backscatter values extracted within a 10 m buffer at the samples' locations.

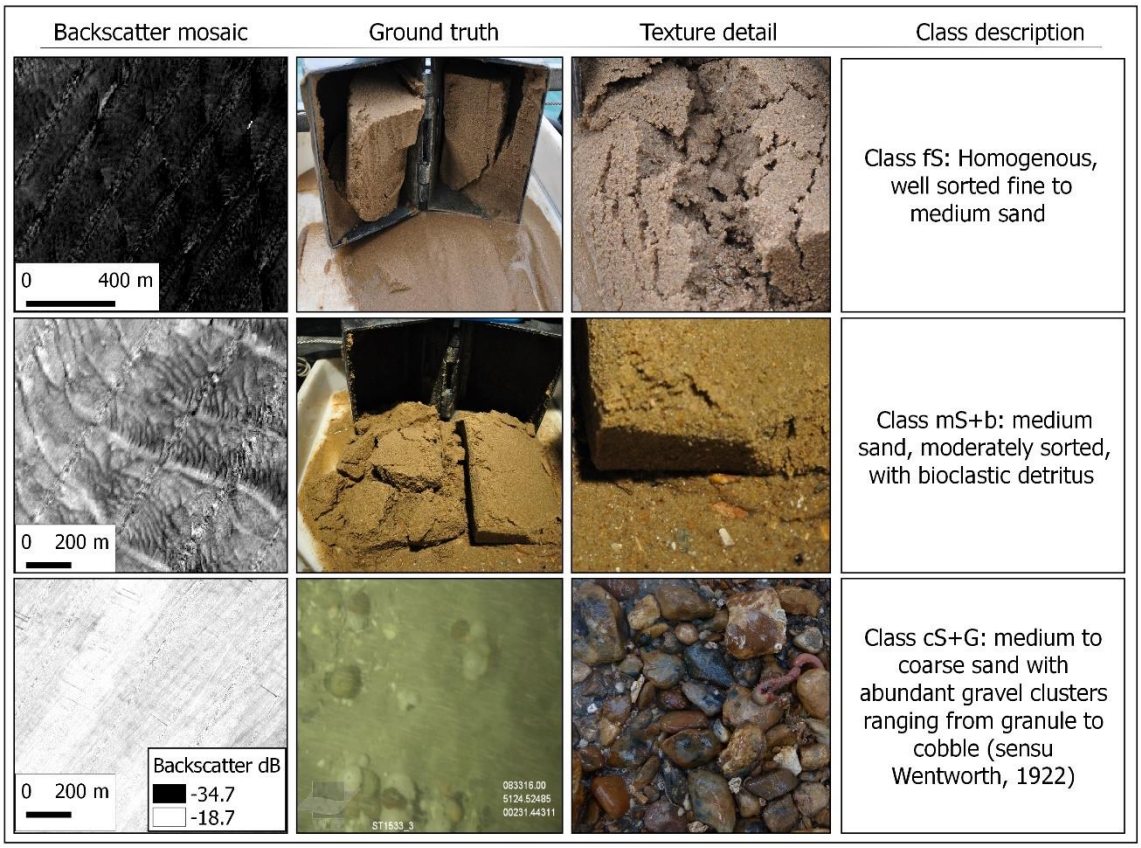


Figure 5.3 – Backscatter mosaic, ground-validation sample picture, textural detail and class description for the identified substrate classes

Table 5.3 - Summary of sample sets used (fS fine homogenous sand, mS + b fine to medium sand with bioclastic detritus, cS + G medium to coarse sand with gravel clusters, VV Van Veen grab sampler)

Class_ID	Training	Test	N-Samples	Gear
fS	9	5	14	Grab (VV)
mS+b	4	2	6	Grab (VV)
cS+G	6	4	10	Video frame
Total	19	11	30	

5.4.5 Morphological evolution

At first, an assessment of the spatio-temporal morphological evolution is carried out to determine whether changes in substrate are due to morphological evolution (i.e. migrating dunes), to an actual reconfiguration of the substrate delineations or to a combination of both. Regions of interest (ROI) encompassing the main morphological and substrate features of the study area were selected to extract 2D profiles from the time series (see Fig. 5.4 for profile locations). Simple yes/no and quantitative metrics of change with information about the directionality (i.e. ebb or

flood dominated bedforms) of the migration can be derived from here. For ease of interpretation, data from 2004 to 2015 were used only (T1 and T7, Table 5.2).

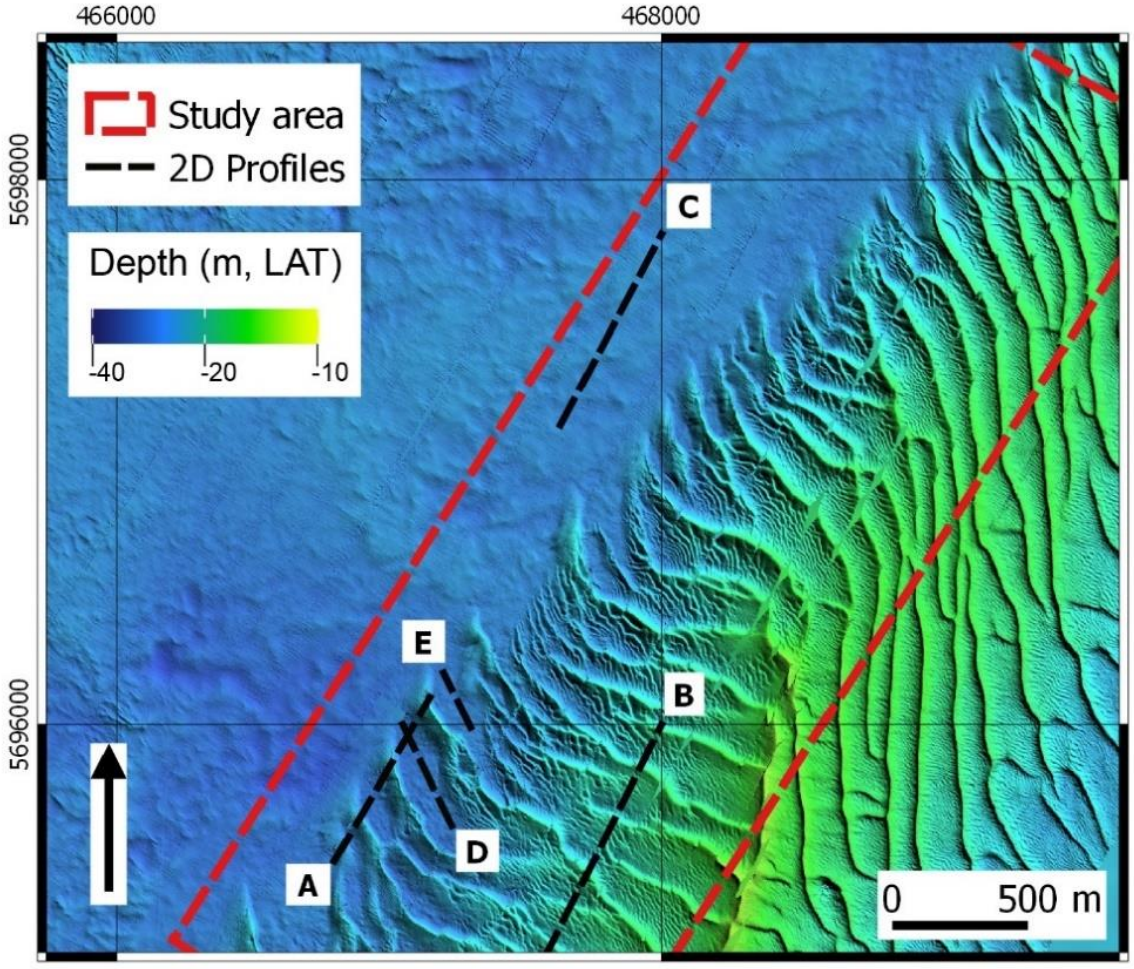


Figure 5.4 - Location of the 2D profiles selected for the analysis of morphological evolution

5.4.6 Supervised classification

The second phase of the analysis makes use of the most recent (T7) acoustic survey for which complementary ground-truth data are available. In order to efficiently combine the two datasets, a supervised classification algorithm is used. Unlike an unsupervised method, where no a priori information about the class labels is provided to the algorithm (i.e. clustering procedures), supervised classification uses ground-truth information to train and test the classification results. The Random Forest (RF; Breiman 2001) algorithm was used for classification. RF has high predictive accuracy in studies focusing on the comparison of supervised classifications of MBES data (Diesing et al. 2014; Diesing and Stephens 2015) and have proven highly successful in data mining research (Li et al. 2016). As explained in Diesing et al. (2014) and Li et al. (2016), the main underlying assumption of this method is that the predictive power of multiple classification trees (the elemental unit of machine learning methods) is higher than that of a single tree. Bootstrapped

samples from the training data are used to construct the individual trees in the forest introducing the first element of randomness. In turn, a random subset of the predictor features is used at the node splits throughout the construction of the model. The result is the construction of unique trees. Decisions about the class allocation (labelling) are made on the basis of majority votes of individual trees. After a feature selection procedure, the RF was run growing 501 trees and leaving the parameters as default. The routine was implemented in R (R Development Core Team 2008) using the RandomForest package (Liaw and Wiener 2002).

5.4.7 Feature selection

A set of textural and morphometric predictor layers were computed from multibeam depth and backscatter grids (Table 5.4). Predictor layers are a set of variables (in this analysis terrain and texture attributes) derived from the MBES backscatter and bathymetry which are combined to the observed substrate type points (response variable) to predict the full-coverage seafloor map (Lecours et al. 2016). The relevance of the predictors was investigated by following the feature selection procedure provided by Kursa and Rudnicki (2010) using the Boruta RF wrapper function. Boruta identifies important variables by performing multiple runs of the RF classification (a total of 1000 runs were performed here) and by comparing the RF Z-scores of the original variables with the scores of their permuted copies (shadow variables). The Z-score is a measure expressing how many standard deviations a score stands from the mean. Higher importance is attributed when the mean Z-score of a variable after n runs is significantly higher than z-scores produced by the shadow variables.

Table 5.4 - Predictor variables dataset with their description

Layer	Description	Type	Software
Backscatter Strength (BS - dB)	256 Grey Level (NG) dynamic range layer – the level of the acoustic energy resulting from the scattering back to its source of emission. Measured as the ratio of the acoustic energy sent and returned from the seafloor, referenced at 1-m from the target at a given incidence angle range.	MBES recorded seafloor backscatter strength	FMGT - QPS
Bathymetry (m)	Post-processed depth samples	MBES recorded seafloor depth	QIMERA - QPS
Roughness (from Depth)	Difference between min. and max. of a cell and its 8 neighbours	Secondary morphometric derivative	Rx64 3.2.3 (raster pkg) by Hijmans and van Ehtten, 2014
Contrast (BS)	Differences of the intensities of the instances within an image in a neighbourhood	Secondary backscatter texture	Rx64 3.2.3 (GLCM pkg) by Zvoleff, 2015
Mean (BS)	Mean filter	Secondary filtered backscatter	Rx64 3.2.3 (GLCM pkg)
Dissimilarity (BS)	Degree of dissimilarity (Euclidean) in a neighbourhood	Secondary texture	Rx64 3.2.3 (GLCM pkg)
Moran (BS)	Spatial auto-correlation in a neighbourhood	Secondary texture	Rx64 3.2.3 (raster pkg)

Moran (from Depth)	Spatial auto-correlation in a neighbourhood	Secondary texture	Rx64 3.2.3 (raster pkg)
Entropy (BS)	Measure of spatial disorder in the distribution of elements within the Grey Level Co-Occurrence Matrix	Secondary texture	Rx64 3.2.3 (GLCM pkg)

5.4.8 Model evaluation

Overall accuracy (A) and Kappa (K) accuracy metrics were derived using the contingency table which cross-tabulates test and predicted instances (Foody 2004). Global accuracy provides a metric expressing the amount of correctly labelled pixels by the classifier whereas Cohen's Kappa reflects the difference between the overall agreement and the agreement expected by chance.

5.4.9 Comparison of thematic maps

Since the supervised information is to be extended to the broader time series of acoustic data for which there is no ground-validation data, an analysis similar to that of Ierodiaconou et al. (2005), in which supervised and unsupervised classifications are compared and evaluated for similarity, was applied. In this paper, K means was chosen as an unsupervised classification method due to its success in finding optimal clustering solutions and after comparing the RF classification to an array of unsupervised classifiers. Hartigan and Wong (1979) algorithm was implemented using the R base functions (R Development Core Team 2008). Given a certain number of classes, the method seeks to reduce and maximise the within and between classes variance respectively by iteratively grouping similar points in their feature space. To validate the application of an unsupervised classifier, paired-pixel metrics of map agreements were computed after Foody (2004), Pontius and Millones (2011), and Pontius and Santacruz (2014). The R package diffeR was used (Pontius and Santacruz 2015). Components of allocation are used to derive the agreement between maps at the level of the entire landscape and per category. Quantity and Allocation describe the amount of change that is respectively due to the proportion of categories between reference and test instances and to the amount of spatial mismatch between categories.

5.4.10 Change detection

Three types of analysis were performed on the backscatter time series in order to extract trends and patterns of change in substrate classes: pre-, post- and ensemble-methods classification. Ensemble approaches combine supervised and unsupervised classifiers, whereas a pre-classification method focuses on the unclassified data values (similarly to directly relying on spectral bands in satellite imagery). The aim of a post-classification approach is to allocate class labels to the data values to produce thematic maps.

5.4.10.1 Pre-classification

The pre-classification approach uses backscatter values taken from rectangular bins of the sampling locations representative of the different geomorphological and substrate features of the ROIs. Following, basic statistics and temporal trends were studied (for example, fluctuations around the ±1 dB accuracy threshold; Hammerstad 2000). In order to detect outliers in the time series, sigma detections where chosen as the favoured statistical measure to quantify the dispersion of a set of data values.

5.4.10.2 Ensemble approach classification

An ensemble method, combining supervised and unsupervised classifications was also applied. K-means classes (dB ranges) identified in T^7 were used to reclassify the complete dataset for which ground-truth data were not available. From this classified dataset, proportion counts were extracted to observe temporal trends. Prior to transforming the successional backscatter mosaics into classified data, the Within Group Sum of Squared Distances plot was computed independently for each dataset. This ensured that the number of classes in each time series was maintained and allowed testing. This also serves to test the class discrimination potential of data gathered at 100 and 300 kHz from the EM1002S and EM3002D, respectively. This technique is similar to computing a silhouette plot where the optimal number and size of classes in a dataset becomes visible (Eleftherakis et al. 2012).

5.4.10.3 Post-classification

The post-classification approach made use of the most commonly employed tool in change detection used in remote sensing studies: the transition matrix (Pontius et al. 2004; Braimoh 2006; Rattray et al. 2013). In this analysis, two unsupervised seafloor maps (e.g. prior and after a natural or anthropogenic event) are cross tabulated to derive detailed statistics describing the temporal changes. Persistence and class swap dynamics, gross gains and losses, between time and between classes' transitions, as well as persistence ratios expressing the tendency of a category to undergo a certain change process were derived after Braimoh (2006). Swap is defined as the change in spatial location of a substrate type between times. The net change describes the difference in quantity of a substrate class between time 1 and time 2. Thus, swap describes changes in location, whereas net change reports changes in quantity. Gain and Loss describe an increase and decrease of the areal extent of a substrate class respectively. Gain (Gp), Loss (Lp) and Net (Np) to persistence ratios are derived as a measure of class tendency to the different types of transition. Values above 1 indicate that a class is more likely to gain or lose from other classes rather than persisting across the time scale analysed. Values close to 0 indicate little or absence of change. The net change to persistence ratio, Np, indicates the overall trend of a category with negative and positive values indicating the directionality of the temporal trends.

5.5 Results

Firstly, results are presented on the supervised classification achieved by implementing the Random Forest algorithm. Secondly, the supervised model is compared to the map of the study area produced by the unsupervised clustering method. Next, the results from the two-dimensional morphological analysis are provided, followed by the change detection approaches tested on the time-series backscatter dataset.

5.5.1 Supervised map of the study area

Figure 5.5 shows the visual validation of the sample datasets. This showed an overall good representation of the BS variation in the study area (Fig. 5.5A). Mean backscatter values, extracted within 10 m circular buffers at the sample's positions, indicate good separation of the classes (Fig. 5.5B) where coarse-hard and fine-soft classes exhibit the highest and lower backscatter values respectively. Similarly, the separation using the bathymetry evidences the distribution of substrate types within different depth zones (i.e. fS on the top of the sandbank, mS + b transiting to the deepest area, and cS + G in the gully; Fig. 5.5C, D). The predicted substrate classes' distribution by the Random Forest supervised classification is shown in Fig. 5.6B. The most important variables selected by the feature selection tool were BS, BS 3 × 3 mean filter, BS Local Moran and bathymetry. With these selected features, the map produced has an overall accuracy (A) of 81%. Furthermore, more than 70% of the classification did not occur by chance (k = 73%).

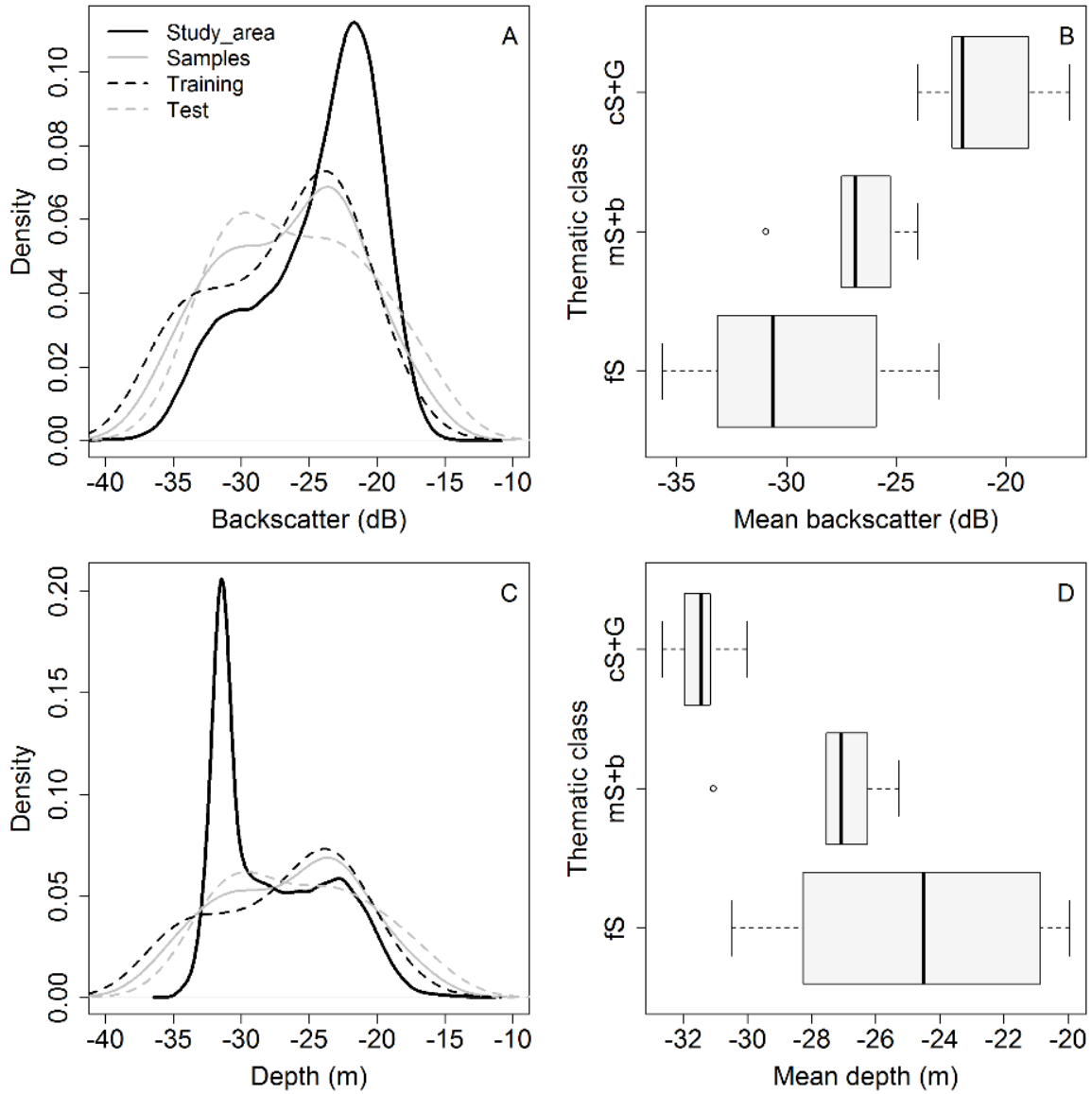


Figure 5.5 - **a** Backscatter distribution in the study area, and per sample dataset (ST1533-T7 dataset), **b** boxplot of mean backscatter extracted from a 10 m buffer at the ground-truth locations, **c** same as (**a**) using depth, **d** same as (**b**) using depth. Training and test refer to the distributions of the training and validation sample datasets used in the RF classification

5.5.2 Comparison between supervised and unsupervised models

Figure 5.6 shows the visual agreement between supervised and unsupervised classifications, while agreement metrics between these models are summarised in Table 5.5. Overall, agreement is high with an overall quantity and allocation difference <10%. In terms of quantity, classes differ by an overall of 0.42%. The larger differences result as allocation disagreement of 9.47 and 8.16% for mS + b and cS + G classes respectively. The fS class is by far experiencing the highest between-map agreement (Table 5) with 1.1 and 0.42 differences in allocation and quantity respectively.

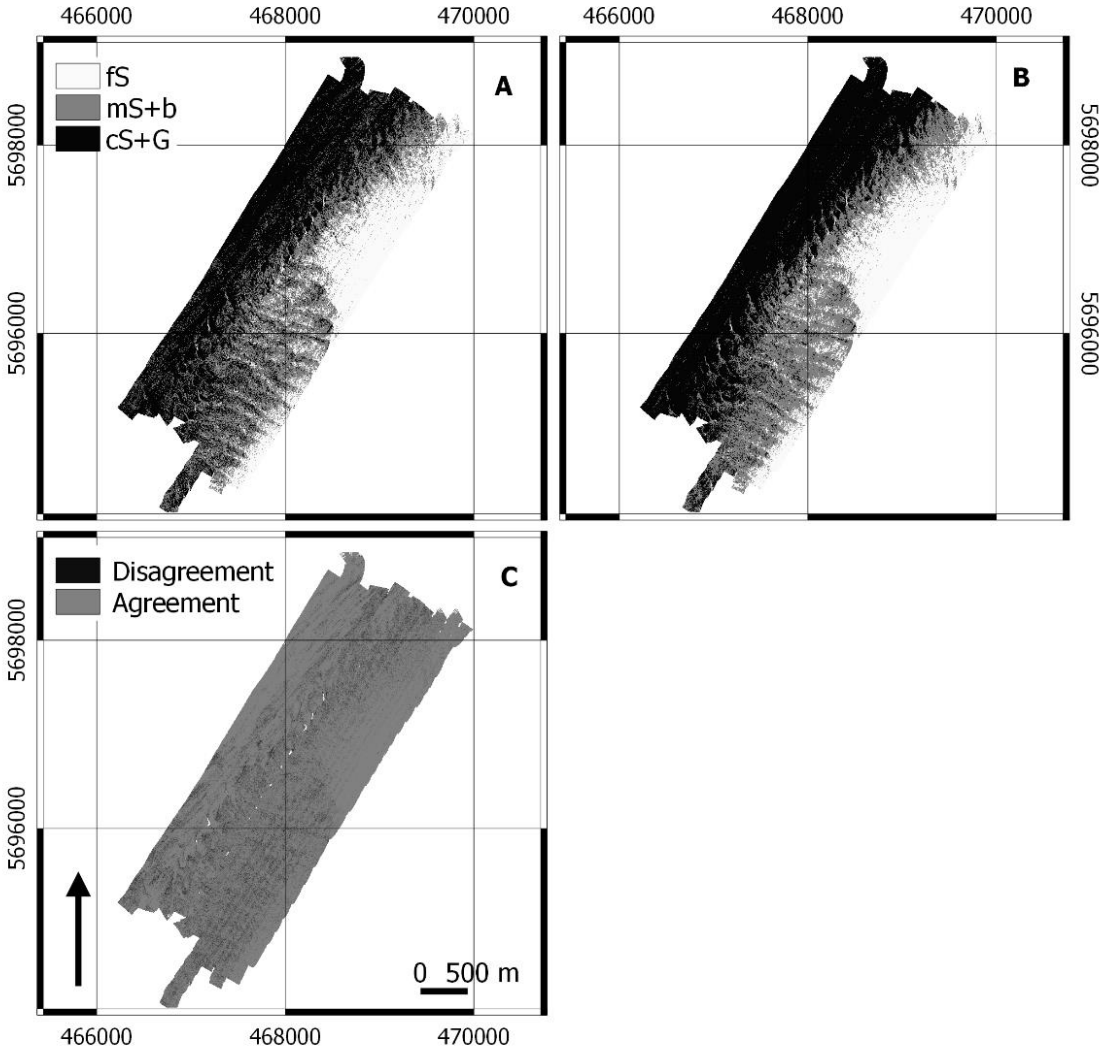


Figure 5.6 - **a** K means unsupervised classification, **b** Random Forest supervised classification and **c** map of overall agreement between classifications

Table 5.5 - Components of difference, allocation and quantity, between models predicted by the Random Forest and K-means (pixels in percentage)

Differences/class	Overall	fS	mS+bio	cS+G
difference	9.79	1.52	9.78	8.28
allocation	9.37	1.1	9.47	8.16
quantity	0.42	0.42	0.31	0.12

5.5.3 Morphological changes

To characterize the dynamics over the full period, depth profiles were extracted from the ROIs for 2004 (T1) and 2015 (T7; Fig. 5.7). Within the barchanoid dunes and along the top sand bank areas (Fig. 5.7A, B respectively) horizontal migration accounts for up to ≈50 m and ≈100 respectively with a SW-NE directionality. Considering the in-between surveys (not shown), it is possible to observe a progressive migration, advancing of ca. 20 m from 2004 to 2010, ca. 10 m from 2010

to 2013 and less than 5 m progressively throughout the remaining surveys up until late 2015. Within the relatively flat and gravel populated areas (Fig. 5.7C, D, F), devoid of dunes, the seabed shows an overall stability. In these areas, vertical changes or aggradation was observed, but cannot be confirmed as they remain within the IHO Order S and 1A confidence envelopes. Nonetheless, a loss of profile complexity is observed between the two campaigns.

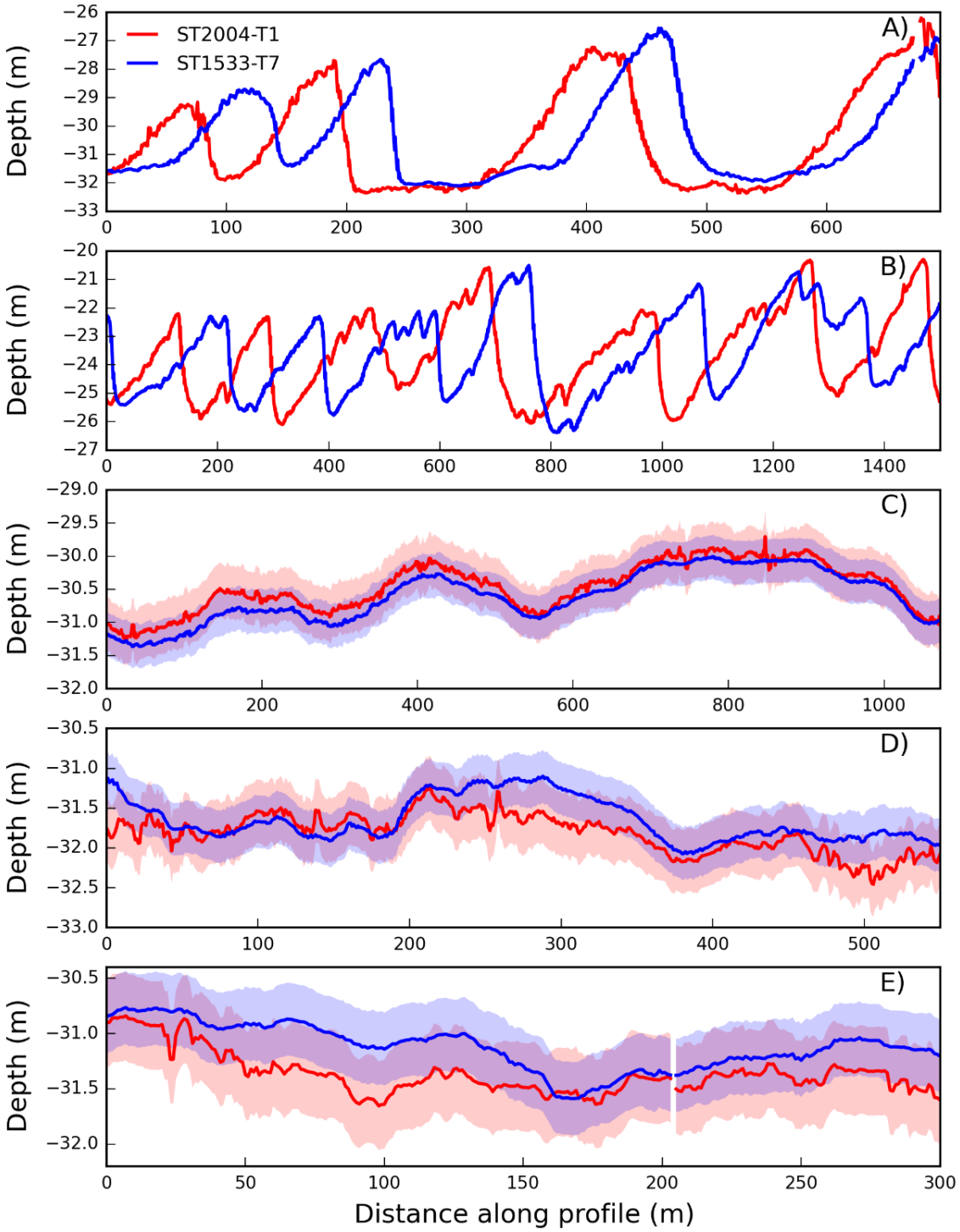


Figure 5.7 - Depth profiles extracted from the digital elevation model time series. **a** Barchanoid dunes area, **b** top sand bank, **c** gully area, **d** gravel refugium 2, and **e** gravel refugium 3. Blue and red envelopes in **d, e:** ± IHO confidence intervals for the EM1002S and EM3002D surveys respectively

5.5.4 Change detection

5.5.4.1 Pre-classification

The boxplot analysis carried out by extracting backscatter data from the selected ROIs is shown in Fig. 5.8. Excluding the EM1002S data (not comparable in terms of insonification values), no significant trends are observable with the exception of zones A and C (transitional and gully zones of the study area) which exhibit deviations >1 σ and generally a decreasing trend up until late 2014 (T5). Noticeably, all selected regions follow an overall elliptical trend (visible in Fig. 5.8H) and re-establish to the initial state of February 2010 (T2) by December 2015. Throughout all cases, the spread is lower than 1 dB evidencing no statistically significant changes. Testing this hypothesis, the reduced $\chi 2$ test computed within each region shows that a significantly negative trend in backscatter spatio-temporal behaviour does not exist ($\chi 2 \ll 1$).

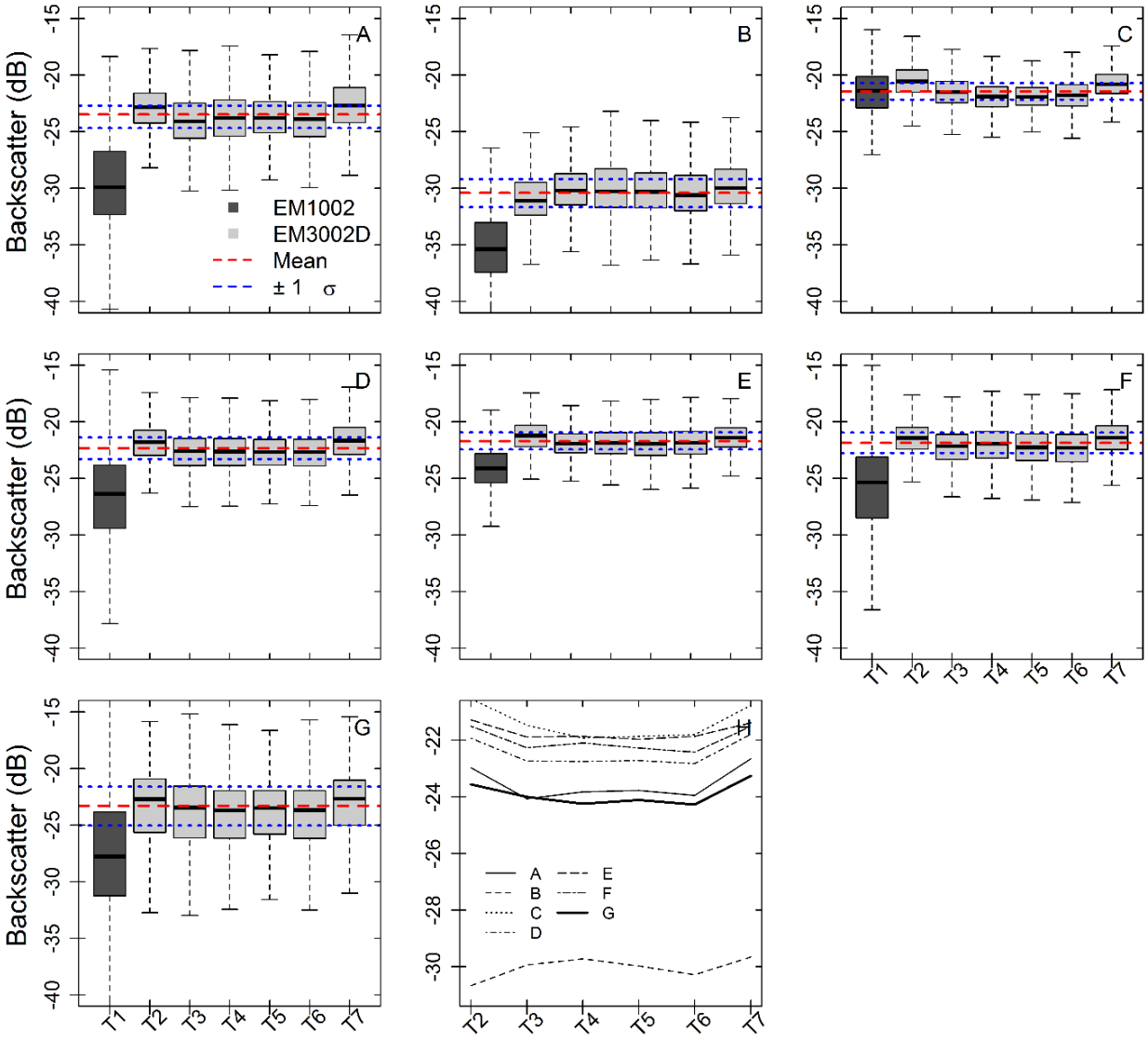


Figure 5.8 - Boxplot analysis for the entire time series (T1–T7). Mean and standard deviation values were calculated from the EM3002D dataset only. **a** Barchanoid dunes area, **b** top sand bank, **c** gully area, **d** gravel refugium 1, **e** gravel refugium 3, **f** gravel refugium 4, **g** entire study area, and **h** mean backscatter values for the EM3002 time series (T2–T7), within each ROI. For **a**–**g** red and blue dotted lines represent weighted mean and ±1 σ error respectively. For the ROIs location the reader is referred to Fig. 1 (A–F boxes)

5.5.4.2 Ensemble approach classification

Class proportion counts per survey were extracted from the classified EM3002D dataset (ensemble approach) and are shown in Fig. 5.9. Temporal trends' and classes' relationships are shown for the entire study area, as well as for the three selected gravel refugia. The fS class appears to be relatively stable across all instances and survey. An inversely correlated relationship is evident between cS + G and mS + b classes. This is also shown in Fig. 5.10 where the proportion counts per survey are plotted against each other. At the level of the entire study area (Fig. 5.9D), and similarly to the pre-classification analysis, this method indicates that the

class proportions return to their original state. On the contrary, within the gravel refugia zones (Fig. 5.9A-C), the cS + G class experiences a net loss in favour of finer substrate types with no indication of re-establishment to a previous state.

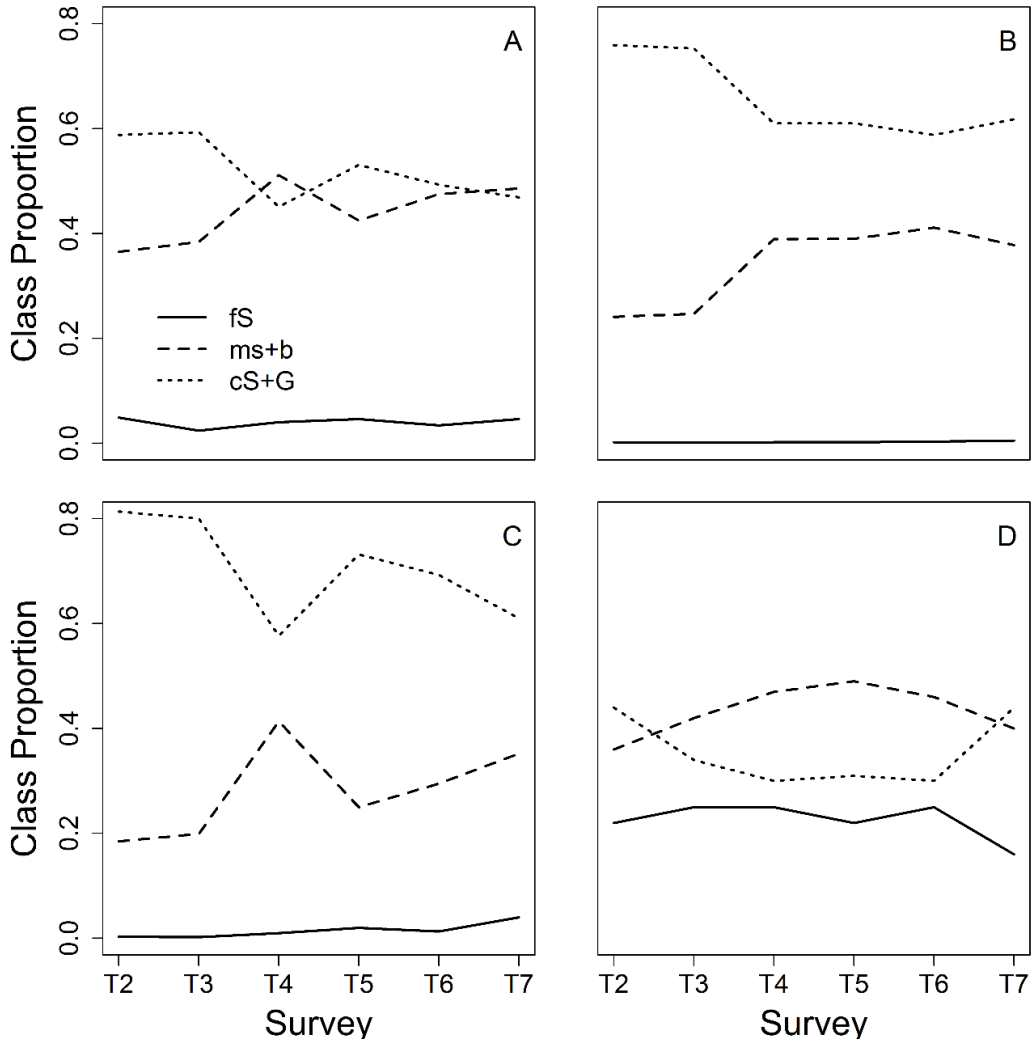


Figure 5.9 - Class proportions during each survey extracted from the classified dataset for three gravel refugia stations (**a**–**c**) and the entire study area (**d**). For the refugia's location the reader is referred to points D, E and F in Fig. 5.1

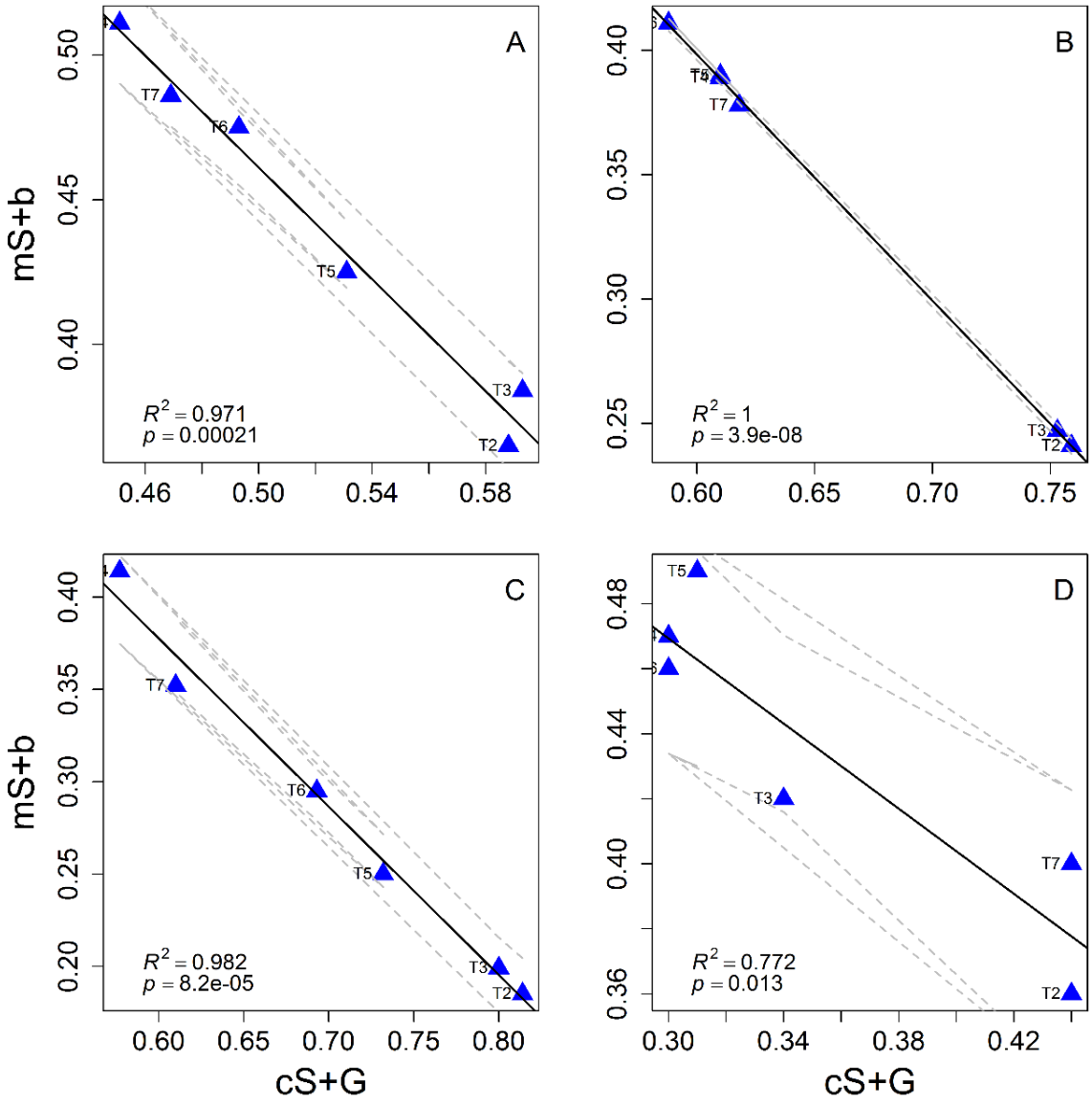


Figure 5.10 - Linear regressions between proportions of cS + G and mS + b classes for the gravel refugia stations (**a**–**c**) and the entire study area (**d**). Dotted lines 95% confidence limits. For the refugia's location the reader is referred to points D, E and F in Fig. 5.1

5.5.4.3 Post-classification

The bi-temporal transition matrix, cross-tabulating the relationships between thematic instances present within the classified maps of 2004 (T^1) and 2014 (T^4) is presented in Table 5.6. Persistence is denoted along the diagonal whereas off-diagonal entries are from-to transitions. Over 50% of the substrate remains static between the classifications. This is mainly driven by persistence of the mS+b and cS+G classes (with 27 and 20% persistence respectively). The class fS experienced the lowest persistence (7.6%) evidencing mostly the dynamics of the bedforms (see Fig. 5.11A where gains and losses result from the migration of dune crests). Following a more detailed inspection of the matrix, ratios describing class tendencies to persistence, gains and losses, swap and net change dynamics were computed

(Table 5.7). Gains, losses and persistence changes are illustrated in Fig. 5.11 where their reciprocal relationships are observable; in particular between mS+b and cS+G classes in the North-Eastern part of the study area (see Fig. 5.11B where the mS+b class gains in favour of the cS+G class, forming ripple marks). All classes experienced a net gain in quantity between the 2 years except for the cS+G class which experienced a net loss of 7.5% [see Fig. 5.11C where it is visible that within a selected refugium, the loss is depicted, partly due to bedform migration (SW) and partly due to the appearance of the mS+b class within the flat and gravelly portion of this area (NE)]. Subtracting the net change from the total change derives swap dynamics. Of the total change for all classes, 83% results as swap; losses in a substrate class are replaced by gains in another substrate class. The mS+b class experienced the highest gain (21.41%), as well as the greatest loss (21.14%) implying that most of the change attributable to this class is due to swap in location. Proportionally, 99.3, 72.5 and 65.3% of changes are attributable to swap for mS+b, cS+G and fS classes respectively. The gain, loss and net changes are compared to the Persistence (diagonal elements of Table 6; calculated after Braimoh 2006) in order to derive ratios (respectively Gp, Lp and Np) providing a measure of class tendency to types of transition. Values above 1 indicate that a class is more likely to gain or lose from other classes rather than persisting between classified instances. Values close to 0 indicate little or absence of change. The fS class has the highest Gp value: it has a high tendency to gain. The mS+b class has similar Gp and Lp ratios, evidencing the high percentage of swap in this class. Most striking is the negative Np ratio and the high Lp of the cS+G class.

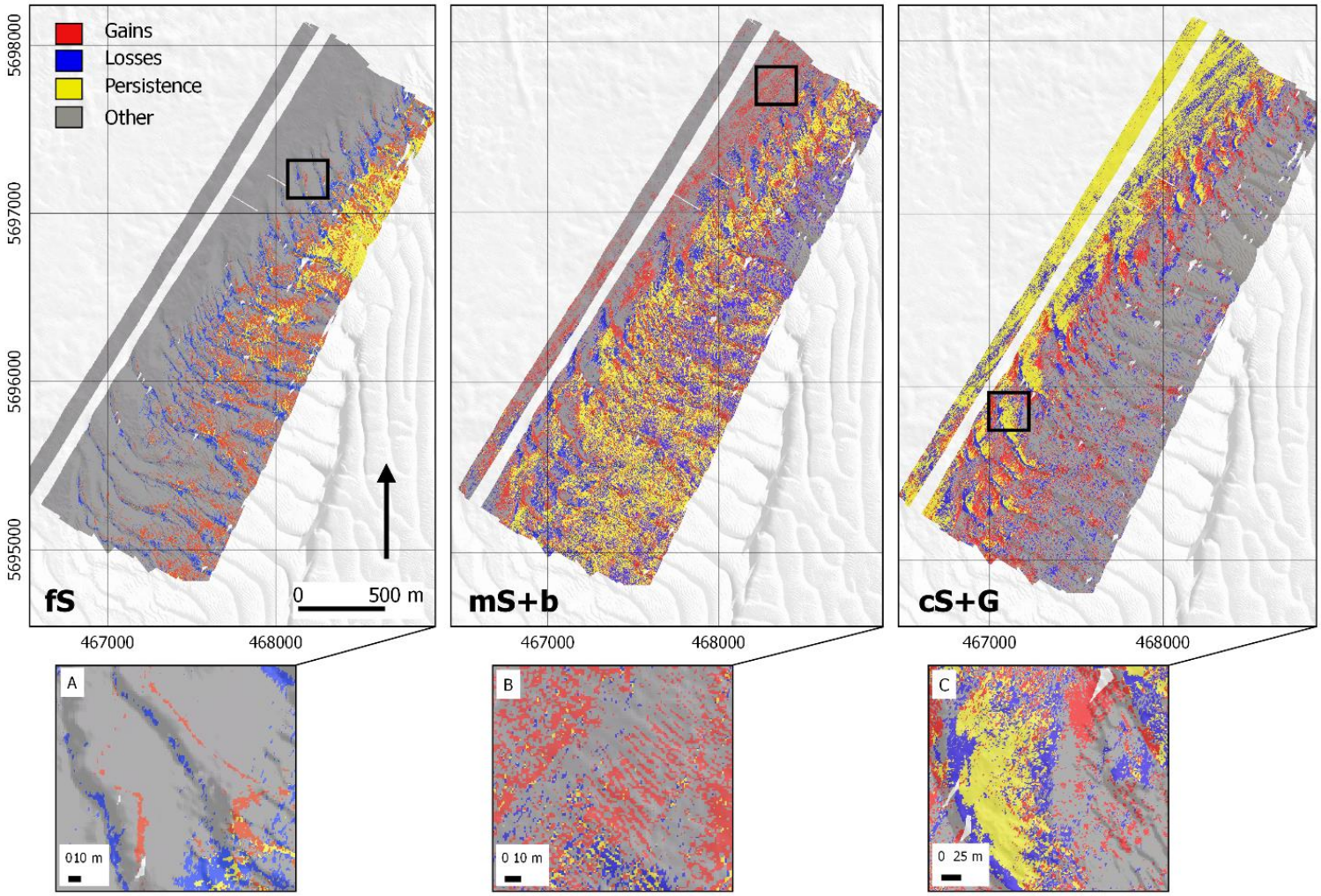


Figure 5.11 - Map representation of persistence, gains and losses for each class in the study area overlapping between T1 and T5.

Table 5.6 - Raw Confusion matrix rounded to two decimals cross tabulating the classified instances in 2004 and 2014 thematic maps

a	fS	mS+b	cS+G
fS	**7.63**	5.89	1.05
mS+b	12.22	**26.79**	8.92
cS+G	1.99	15.53	**20.03**

Table 5.7 - Summary of the changes between 2004 and 2014 (in percentage and expressed as ratios)

Class	Total 2014	Total 2004	Gain	Loss	Total change	Net (Quantity)	Swap (Location)	N_p	G_p	L_p
fS	21.83	14.56	14.21	6.94	21.15	7.27	13.87	0.96	1.87	0.91
mS+b	48.2	47.92	21.41	21.14	42.55	0.28	42.28	0.02	0.8	0.79
cS+g	29.99	37.53	9.97	17.51	27.48	-7.55	19.93	-0.38	0.5	0.88
Total	100.02	100	45.59	45.59	91.18	15.1	76.08			

5.6 Discussion

5.6.1 Multibeam backscatter in a monitoring context

The basic premise in using MBES backscatter data for seafloor change detection is that changes in substrate cover must result in changes in backscatter values and changes in backscatter due to seafloor cover change must be large with respect to changes caused by other factors (readapted from Singh 1989) such as sea conditions, sensor's intrinsic characteristics, changes in on-board acquisition parameters, vessel speed and direction of survey (Rattray et al. 2013; Lurton and Lamarche 2015). As such, verification of MBES backscatter stability is critical and should be controlled (Anderson et al. 2008). In this study, these limitations were mostly overcome as the dataset used was acquired by maintaining rigorous standards of acquisition and processing, including careful attention on environmentally dependent transmission losses (i.e. by regular control of absorption coefficient). To verify instrumental drift on the medium to long term, the trend in backscatter levels was compared against a time series in backscatter levels at a known reference area (KWGS reference area; Blue Polygon in Fig. 1; Roche et al. 2016). As such, average backscatter levels of the RV *Belgica* EM3002D could be compared from one survey to another during a similar period and allowed obtaining a dataset with temporally consistent dB ranges (yet relative). However, changing environmental factors and seabed conditions may affect the backscatter values also. The effect of biological activity, which is seasonally driven and linked to the spawning and recruitment period of benthic species, is probably the most prominent factor. From literature, it is known that megabenthic zoo- and/or phytobenthic structuring species can be responsible for significant changes in the acoustic signal (e.g. Demosponges and Submerged Aquatic Vegetation, brittle stars; Montereale-Gavazzi et al. 2016; Holler et al. 2016), but also the occurrence of soft substrata

macrobenthos ecosystem engineers such as tubeworms and some bivalves (e.g. Degraer et al. 2008; Van Lancker et al. 2012). Hitherto, the impact on the actual backscatter levels is poorly quantified and more research is needed on this aspect in a monitoring context. Beside changes due to the successional stages of some benthic species, natural variability in sediment deposition and erosion can also affect the backscatter level. This will depend on the hydrodynamics of an area, as well as on the sediment availability. Collection of tightly spaced acoustic surveys would be ideal to have a better control on the driving forces which would support the interpretation of trends in backscatter levels. In this study, the time lag between surveys was rather irregular which complicated distinguishing changes from natural versus anthropogenically-steered events. The combination of morphological analyses with backscatter change analyses is important in this regard.

5.6.2 Change detection methods

The pre-classification analysis of the backscatter time series indicated that within the selected regions of interest, no significant changes in seabed substrate could be detected across the timespan analysed. Since the first dataset was recorded with a former-generation echosounder, which was not cross-calibrated with the EM3002D and using a different frequency range, the values derived could not be directly compared in terms of the range in insonification values. The only evident behaviour in the data was in the barchanoid and gravel gully regions where locally, the mean backscatter level fluctuated around the 1 σ deviation (Fig. 8a, c). Since the comparison is rather focused on the spatial delineation and areal extent of the substrate classes rather than the intrinsic, physical characteristics of a circumscribed area, the post-classification methods, as adopted similarly to Rattray et al. (2013), did allow comparing data from different echosounders, and acquisition parameters. This was also possible due to the agreement in the number and size of classes discriminated by the two echosounders. The approach revealed information on the behavioural tendencies of certain substrates to undergo a certain change such as the negative trend of the hard substrate class and gain of the finer substrates. In this study, there was a high agreement between supervised and unsupervised models, using quantity and allocation agreement/disagreement metrics, which allowed extending the analysis to the entire time-series dataset. As such, the initially ground-validated information could be fully exploited and extended to the full backscatter dataset. Substrate class proportions over time could be extracted so that global changes could also be accounted for. This is unlike the pre-classification approach that is limited to selected sub-areas where backscatter levels were extracted from. Therefore, the ensemble approach combined supervised and unsupervised classification algorithms (similarly to Ierodiaconou et al. 2005) which allowed using one ground-truth dataset to train a classification that was subsequently applied to the whole time series. This is a big advantage since sampling of each time series is most often not realistic given survey time and cost restrictions. Here, application of consistent data acquisition and processing allowed the comparability of instant

statistical analysis results at various times. According to recent research (Li et al. 2016), it was shown that the Random Forest classifier is a highly valuable tool for seafloor applications, producing accurate models and providing information on the most important feature layers used in the classification. Similarly to Diesing et al. (2014), backscatter was by far the most important variable for seafloor substrate discrimination (with highest Boruta Z-score after 1000 runs). Depending on the method applied, the accuracy of the change detection is strictly dependent on the accuracy of the classified maps used in the assessment and on the stability of the repeated observations.

5.6.3 Application within a MSFD context

By classifying the data, it was shown that from before the start of dredging activities (T1), northwards of the study area, to just after the peak of marine aggregate extraction (T5), the gravel class progressively decreased at the level of the entire study area, including a net loss of the gravel class extent within the defined ecologically noteworthy areas (Fig. 1, black outlined polygons). From this, the ratio of hard versus soft substrata (Belgian MSFD indicator on seafloor integrity) first showed a negative trend, at least after the peak of the extraction activity, followed by a positive trend indicating a recovery process. Based on the depth time series, a morphological analysis revealed that part of the change is attributable to bedform migration of which the drivers require further investigation. An aggradational trend in the gravel areas was suggested from the observations, though this fell within the IHO confidence limits used. Despite this, changes in the depth profile depicted a reduction in seafloor complexity considering the surveys before and after the initiation of intense marine aggregate extraction. A methodological framework to unambiguously link changes to pressures is under development and is yet hampered by a lack of data and knowledge on the natural variability and resilience of offshore sedimentary systems. Nevertheless, the present results are highly significant from an ecological perspective and necessitate a further investigation of the substrate evolution. If indeed smothering and/or deposition events would be more persistent under increased anthropogenic pressure, this may affect several ecosystem states and functions: e.g. reduction of sessile bio-encrusting epifauna; loss of surficial complexity leading to reduced micro-roughness; burial of biogenic clastic material; and overall reduced potential of benthopelagic coupling (Watling and Norse 1998; Hewitt et al. 2005).

5.7 Conclusion

This study highlights the importance of researching approaches and testing tools usable for local- and regional scale environmental assessments (i.e. for MSFD implementation). A selection of useful methodologies was presented to detect changes in seafloor substrate types. The investigation showed how under specific standardised multibeam backscatter acquisition procedures, the confidence of

repeated acoustic observations could be enhanced significantly and how the valuable, but expensive ground truth information could be propagated from one survey to a time-series dataset via the application of supervised and unsupervised classification routines. The serial backscatter dataset was analysed using techniques developed in the remote sensing terrestrial realm showing that the methodologies are applicable for marine environmental monitoring. This is most promising for before and after control impact (BACI) type of assessments and such datasets would inevitably increase our understanding of anthropic impacts over an area. Although the methods presented were tested at local scales, they are repeatable and can be applied to broad-scale geographical extents; a major limitation being the need to collect large-scale datasets covering entire jurisdictional areas.

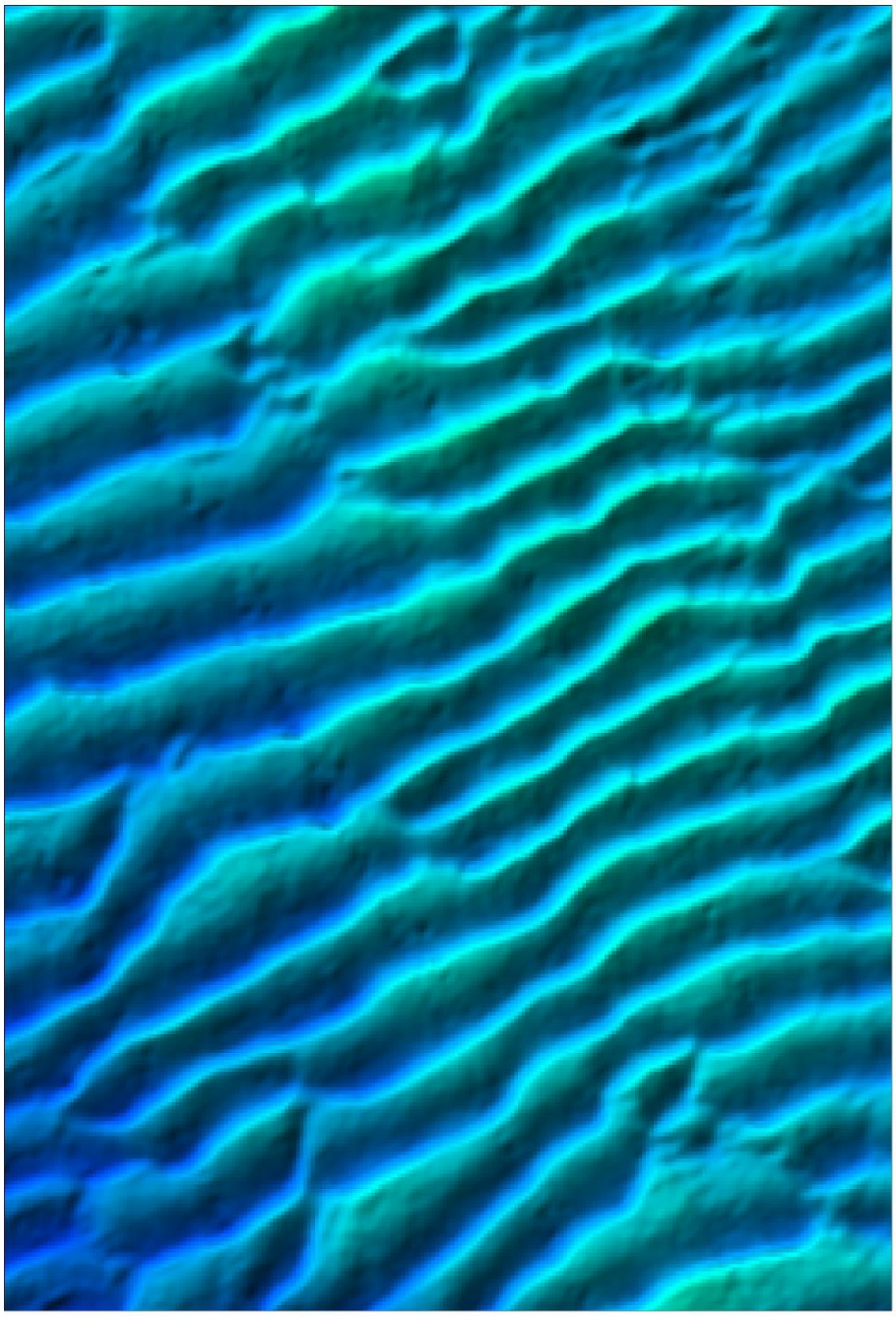
Acknowledgements: This paper is a contribution to the projects MOZ4 (Flemish Authorities, Agency Maritime Services and Coast, Coast, contract 211.177), INDI67 (BELSPO, contract BR/143/A2/ INDI67), ZAGRI, a federal Belgian programme for the continuous monitoring of sand and gravel extraction, paid from private revenues, and TILES (BELSPO, contract BR/121/A2/TILES). Ship time RV *Belgica* was provided by BELSPO and RBINS–OD Nature. The Renard Centre of Marine Geology of Ghent University (RCMG) is acknowledged for the use of the sediment analysis facilities and instruments. Flanders Marine Institute (VLIZ) is acknowledged for the use of the video frame. Flanders Hydrography is acknowledged or the use of background bathymetric data. Two anonymous reviewers are acknowledged for their useful comments and improvements on the initial manuscript.

Author Contributions: Conceptualisation, G.M.-G, V.V.L.; Formal analysis, G.M.-G., N.T; Investigation, G.M.-G., V.V.L; Methodology, G.M.-G., M.R., K.D., and N.T; Project coordination, V.V.L., M.R. and X.L.; Writing—Original Draft, G.M.-G.; Writing—Review and Editing, G.M.-G., N.T., M.R., X.L. and V.V.L.

Addendum – *Errata corrige*

The interpretation of the morphological evolution presented in Page 239, Figure 5.7 has been corrected. Only values for Fig. 5.7A where reported and those of 5.7B omitted. Therefore, the 100 m migration over the full period of 11.5 years was specified in the interpretation. Furthermore, a mistake at page 227 the expression “well stratified” was corrected to “well mixed”.

Chapter 6



Sand ripples from DTM. Westdiep, BPNS, Belgium.

6. Discussion

6. Discussion

Under the core principles of an ecosystem-based approach, developing an explicit knowledge of the spatio-temporal trends of benthic substrates, and of the methodologies to obtain such information, are fundamental prerequisites for the tutelage of benthic habitats (Long et al., 2015). This requires testing the data from state-of-the-art seafloor mapping and sampling gears in use today and developing methodologies towards the data-integration and the interpretation of complex data-structures. Ultimately, these approaches must be objective, accurate and repeatable as well as operationally feasible and ecologically meaningful.

As our anthropogenic activities interact with the marine ecosystem over broad spatial scales (> 1km), and ecological management requires understanding of meso (10m-1km) to fine (< 10m) spatial scales, the acquisition of adequate seafloor contextual information fulfilling the emerging maritime management issues largely relies on the use of underwater acoustics. Here, multibeam echosounders excel in performance; able to acquire fine scale observations over broad spatial extents (Kenny, 2003). In this regard, the rationale motivating this doctoral thesis is the fact that the benthic substrate is a keystone building block of benthic habitats as it determines suitability for given benthic biota and benthic communities (Diaz et al., 2004; Miller et al. 2002; McArthur et al., 2010), especially in predominantly sedimentary marginal Northern Atlantic continental shelves where foundation species are less frequent. Furthermore, besides measuring its type, extent and status over space, its temporal assessment is a key component of a monitoring programme targeting the quantification of anthropogenically induced seafloor changes (Miller et al., 2002).

In this doctoral thesis a variety of fundamental scientific challenges were identified and assessed critically, enhancing our knowledge on the way we can use innovative approaches to assess the seafloor status, both in space and time. Particularly, this research endeavoured reducing of challenges associated to the operationalisation of multibeam backscatter measurements for substrate/habitat mapping and environmental monitoring objectives set out by the Belgian State (2012) and under the broader context of supporting the implementation of the Marine Strategy Framework Directive (MSFD, 2008/56/EC) (reported in Chapter 1b). To that end, various aspects of seafloor mapping and monitoring based on multibeam technology have been methodologically put forward and tested in three independent (yet related) research chapters (namely Chapters 3, 4 and 5).

The foundation of the hydroacoustic measurements, ground-truthing and classification approaches as used in this thesis, is the advanced characterisation of the immediate seafloor substrate nature and the relation of the acoustic measurements to the physical and measurable characteristics of the seafloor. While seafloor substrate mapping is not a novel practice to marine sedimentology and

geology, it is the novel application of state-of-the-art scanning sensors combined with novel approaches in automated and objective data-integration that drove the establishment of Acoustic Seafloor Classification as a scientific multi- and trans-disciplinary field in its own right (Anderson et al. 2008), ramifying into a range of disciplines and benefitting in particular modern approaches to marine seascape ecology and benthic habitat mapping (Brown et al., 2012). The improvement of the ecological meaningfulness of spatio-temporal patterns inferred from the current generation of surveying and sampling instrumentation (i.e. how do detected changes relate to ecological processes of interest in view of monitoring seafloor integrity), is undoubtedly the forthcoming main scientific objective of these evolving approaches.

Based on the recognition that the efficiency of modern marine management approaches is hindered by a paucity of objective and repeatable methodologies to gain accurate information of the seafloor nature both in space and time, and that there are several steps and scientific challenges in setting-up a seafloor monitoring strategy, a set of core research objectives were identified in Chapter 1b and can be broadly summarised under the following three steps (related to points A, B and C in Chapter 1b) where underneath each, the specific objectives dealt with in this thesis are reiterated:

I. Recognising the challenging operational environment of this part of the North Sea, the first step concerned improving on multibeam surveying and ground-truthing as well as processing strategies, to establish a baseline MSFD survey effort for the future monitoring of seabed substrates at both site-specific and regional levels.

Objectives:

- *To test supervised and unsupervised MBES and ground-truth data integration approaches towards the objective, repeatable and accurate predictive mapping of benthic substrate type: testing the influence of MBES data on the accuracy of predictive models.*

- *To test the MBES backscatter discriminative ability in terms of number and type of differing seafloor classes mappable and investigate sediment-acoustic relationships.*

- *To assess the presence of trade-offs between predictive thematic accuracy and substrate classification schemes employed.*

II. Recognising the advancing potential of absolute calibration and the potential to use multibeam backscatter as a direct proxy of seafloor change, the second step related to quantifying external (environmental) sources of variance causing discrepancies between repeated backscatter surveys. This was approached by designing and conducting dedicated

field-experiments, focusing on the study of short-term (half-diel time-scale) environmental variability.

Objectives:

- *To critically evaluate and quantify the magnitudes of the potential short-term, tidally-induced environmental sources of variance that could lead to unwanted backscatter signal fluctuations.*

- *To question the implications these may have on the detection of longer-term changes and on the interpretation of patterns in acoustic backscatter imagery.*

- *To assess whether these variability sources can be bypassed and/or corrected where needed.*

III. In response to the paucity of change detection approaches in the marine mapping literature and preparatory to future MSFD monitoring cycles, the third step focused on testing image-based change detection approaches

Objectives:

- *To test and assess the potential of different change detection methodologies based on serial backscatter measurements.*

- *To evaluate the kind of spatio-temporal patterns observable and quantifiable.*

These objectives were fulfilled by the research accomplished through three related, yet independent chapters (namely Chapters 3, 4 and 5). Together, these chapters represent a novel contribution to both our understanding of the spatially-explicit variation of the benthic substrates in space and time and to the methodologies and approaches used to model and predict them. Therefore, the work contributes to advance the operationalisation of MBES seafloor backscatter measurements and acoustic seafloor classification to monitor the water-sediment interface status in general, and in fulfilment of the Belgian MSFD objectives.

Hereafter, the reiterated steps and objectives are discussed in more detail and critically evaluated under two main themes: (1) *Towards a seafloor mapping strategy for the Belgian Part of the North Sea: setting the MSFD baseline survey effort*; and (2) *Seafloor monitoring using MBES: variability and change detection.*

6.1 Towards a seafloor mapping strategy for the Belgian Part of the North Sea: setting the MSFD baseline survey effort.

Mapping is considered the primary and indispensable step in the context of environmental protection and any monitoring application relies on an initial mapping effort (Pickrill and Kostylev, 2007 and Chapter 2 “Backscatter for discovery”). Due to the challenges associated with full-regional-coverage seafloor mapping, the surveying efforts often comprise the strategic prioritisation of ecologically noteworthy areas (Strong, 2015) in respect to the salient anthropogenic activities that may alter the predetermined seafloor status (see Chapter 3 and Table 6.2). This is particularly the case for the BPNS where issues inherent to obtain regional full-coverage mapping are linked to the overarching number of anthropogenic activities taking place over a remarkably busy and limited spatial extent, with intense vessel routed navigation and frequent inclement weather conditions: together hindering the success of mapping and sampling activities.

To that end, the research presented in Chapter 3 endeavoured in setting up objective and repeatable surveying strategies and hydroacoustic and ground-truth data integration methodologies, producing detailed seafloor substrate maps, unpreceded in accuracy and quality in Belgian substrate mapping studies (Chapter 3 for a comparison with former generation echosounders and processing and classification routines), testing methodologies reproducible farther afield, and providing a thorough baseline mapping effort for future MSFD-oriented monitoring campaigns.

Prior to the data-integration phase, comes the planning and implementation of surveying strategies. In this regard, in Chapter 3 it was shown how a considerable surveying effort (approximately 150 km^2 over ~40 days of navigation during 2015-2018; ~15 days per year; 2-3 missions per year), covering priority areas targeted by site-specific monitoring (fine-scale patterns < 10 m), as well as broader scale surveys targeting a more regional-level monitoring (meso-broad patterns > 1 km), can be obtained with relatively limited ship time. Table 6.1 details some of the key aspects that were kept in mind in setting up the surveying strategy (besides those listed in the “Multibeam survey strategy” paragraph of Chapter 3).

Table 6.2 details the envisaged monitoring targets, reiterating the surveyed locations in the BPNS and reporting their general morphosedimentary and benthic habitat type (*sensu* macrobenthos suitability habitat map – Degraer et al., 2008 - and knowledge of epibenthic communities from pioneering and recent research - Houziaux et al., 2008; 2011) along with the targeted scale of assessment of the monitoring (i.e. fine or broad scale). Further, the table includes details of the salient anthropogenic activity possibly influencing the surveyed area in the near and/or far field, the potential cause of natural variability and the targeted change detection phenomenon (i.e. phenomenon that could be detrimental to benthic biota – Miller et al., 2002) capturable by a combination of serial MBES, ground-truth data and modelling

approaches. The table will be recalled to in the "*Seafloor monitoring using MBES: variability and change detection"* section of the discussion.

Table 6.1 – Surveying factors of importance when acquiring backscatter for monitoring. Table constructed in accordance to the guidelines and recommendations set out in Lurton and Lamarche (2015) and the Seafloor Mapping Multibeam Field Manual by Lucieer et al., (2018)

Survey planning and data acquisition	**Approach**	**Justification**
Backscatter Reference Area Survey (@ BPNS → Roche et al. 2018)	At the launch of each oceanographic mission acquire data over pre-defined survey lines in the designated reference area.	Control measurements stability; Allow for inter-system-calibration-propagation; Potential for absolute calibration; Bathymetric and Backscatter "Patch Tests".
Sound velocity	Accounted by SVP at transducers i.e. (near surface). Note: this approach satisfies the shallow/non-stratified waters of the BPNS. Where this does not apply, CTD casting will be mandatory.	Correct depth conversion.
Control of the absorption coefficient (*sensu* Francois and Garrison, 1982a, b)	Onboard computation of absorption coefficient @ 30 min. intervals. Data from ODAS or MIDAS system on RV *Belgica* and *Simon Stevin* respectively.	Adequate correction of the dissipation of acoustic energy due to hydrological status of the seawater medium (i.e. temperature, salinity and pH). BS Radiometric correction.
Data-logging along isobaths	1) Navigation parallel to isobaths (i.e. not up or downslope)	1) Maintain reasonably constant coverage along survey lines. Navigation along prevalent tidal axis (and seafloor directionality). Moreover, Suitable vessel draught depths for the RVs in use in the BPNS (max 4.8 m).
Survey line overlap	1) Min. 30 % for *box surveys* (full-coverage); 2) Ideally Min. 3X rep. parallel lines for *reconnaissance surveys* (trajectory surveys) and > 30 % overlap.	1) Overlap outer beams. Account for line keeping errors; 2) Allow sufficient overlap for forthcoming repeat surveys along coastal and offshore trajectories.
Speed above ground	Constant surveying speed. Max. 8-9 kt.	Minimise under keel aeration. Well distributed sounding density = coherence in gridded data resolution.
Sea state	Max. Wave-height = 1 m.	Data artefacts = poor quality data. Decreased performance of motion compensation. Under keel aeration. Etc.
Acquisition platform settings	Strictly standard acquisition runtime parameters (Aligned to surveying standards of the Continental Shelf Service f the Federal Public Service of Belgium)	Data comparability. Stability control. Repeatability. Seamless mapping.
Data quality check/Sampling	Onboard BS processing → GIS integration → Samples planning --> Strictly within Max. 48 h from the survey.	Monitor and amend survey coverage. Planning of ground-truth sampling locations respect to acoustic facies in BS CBI.

Table 6.2 – Pt. 1. Summary of envisaged monitoring sites with specifications of: (1) depth zonation and morphosedimentary type, (2) expected benthic habitat type, (3) spatial scale of investigation (in the monitoring framework), (4) near and far field anthropogenic activity, (5) natural drivers of change, (6) potentially mappable/detectable changes using a combination of MBES and ground-truth, (7) potential geo-sedimentological impact, (8) potential bio-ecological impact, and (9) study site legal designation (sensu MSP). Notes are reported in the bottom row.

Survey area (Ref. Fig. 3.1b Ch. 3)	Depth zonation (Max. depth) /Morphosedimentary type	Expected benthic habitat/community	Monitoring assessments level and spatial scale of investigation	Potentially influencing nearest anthropogenic activity	Salient natural drivers of change
MSFD Coastal trajectory	Nearshore, -15 m (mS-S/sM-M)	*Nepthys cirrosa, Ophelia limacina, Macoma baltica, Abra alba*	Regional (meso/broad > 1 km)	BT (NF), DD (NF, FF)	Large and Small-scale bedform migration, regional sediment transport
MSFD Offshore trajectory	Offshore, -30 m (S-gS-sG-G)	*Nepthys cirrosa, Ophelia limacina, Macoma baltica, Abra alba*	Regional (meso/broad > 1 km)	BT (NF)	Large and Small-scale bedform migration, regional sediment transport
Hinder Banks	Offshore Swale, -35 m (gS-sG-G)	*Nepthys cirrosa, Ophelia limacina* (Arborescent Epifaunal and macrobenthic assemblages of the "gravel fields")	Site-specific (fine < 10 m)	BT (NF), MAE (FF)	Large and Small-scale bedform migration, regional sediment transport, Mobile sand (sand patches, dunes or ribbons), gravel lineations
Westdiep	Nearshore Swale, -10 m (S-gS-sG)	*Nepthys cirrosa, Ophelia limacina, Macoma baltica, Abra alba* (*Lanice conchilega* "reef")	Site-specific (fine < 10 m)	BT (NF)	Large and Small-scale bedform migration, regional sediment transport
Oostende disposal ground and MOW 1	Nearshore Flat, -10 m (sM-M)	*Macoma baltica, Abra alba*	Site-specific (fine < 10 m)	BT (NF), DD (NF)	Large and Small-scale bedform migration, regional sediment transport

Kwinte swale	Offshore Swale, -25 m (S-gS-sG)	*Nepthys cirrosa, Ophelia limacina* (Arborescent Epifaunal and macrobenthic assemblages of the " gravel fields")	Site-specific (fine < 10 m)	BT (NF), MAE (NF-ceased)	Large and Small-scale bedform migration, regional sediment transport
Goote Bank swale	Offshore Swale, -25 m (S-gS-sG)	*Nepthys cirrosa, Ophelia limacina* (Arborescent Epifaunal and macrobenthic assemblages of the " gravel fields")	Site-specific (fine < 10 m)	BT (NF), MAE (Exploration)	Large and Small-scale bedform migration, regional sediment transport
Thornton Bank swale	Offshore Swale, -30 m (S-gS-sG)	*Nepthys cirrosa, Ophelia limacina* (Arborescent Epifaunal and macrobenthic assemblages of the " gravel fields")	Site-specific (fine < 10 m)	BT (NF), WE (FF), FF (FF)	Large and Small-scale bedform migration, regional sediment transport
Northern exploration zone	Offshore Swale (plane), -45 m (S-gS-sG)	*Nepthys cirrosa, Ophelia limacina* (Arborescent Epifaunal and macrobenthic assemblages of the " gravel fields")	Site-specific (fine < 10 m)	BT (NF), unknown	Large and Small-scale bedform migration, regional sediment transport, Scouring
*Note	Nearshore = within 12 nautical miles	(Degraer, 1999; Houziaux et al., 2008; Van Hoey et al., 2004)	*Altogether the surveys are for broad-scale assessments though individually fine-scale assessments can be made. Note that the mapping unit remains fine-scale (1 to 5 m pixel size)	NF = Near Field, FF = Far Field, BT = Bottom Trawling, WE = Wind Energy, MAE = Marine Aggregate Extraction, DD = Dredging and Dumping, FF = Fish Farm	Storms and wave action to be considered at all study areas

Table 6.3 – Continued.

Survey area (Ref. Fig. 3.1b Ch. 3)	Targeted areal and physical phenomenon in change detection based on MBES and GT and in respect to IA or Baseline mapping effort	Impact on seafloor morphosedimentary aspect	Implication of impact for benthic biota	Environmental/Legal designation

MSFD Coastal trajectory	Spatial reorganisation of main granulometric classes. S encroachment (sandification). Westward extension of sM to M	Erosion, Remobilisation, Siltation, Scouring (Lineations)	Reduced habitat suitability, suffocation, complexity, anoxia	SW part crossing Habitat Directive, Special zone for seabird protection
MSFD Offshore trajectory		Erosion, Remobilisation, Scouring (Lineations)	suffocation, anoxia, removal, crashing	Habitat Directive Area
Hinder Banks	Monitoring of the sand/gravel ratio, Seascape changes (patch level)	Erosion, Remobilisation, Scouring (Lineations), Smothering, Fine sediment entrainment, Overtopping	Smothering, Abrasion (loss of substrate/complexity)	Habitat Directive Area, Special zone for seabed integrity
Westdiep	seascape changes	Erosion, Scouring (Lineations)	Abrasion (loss of substrate/complexity)	Habitat Directive Area
Oostende disposal ground and MOW 1	Spatial reorganisation of main granulometric classes. S encroachment (sandification). Westward extension of sM to M, Establishment of "disposal mounds"	All of the above	Smothering (crashing, suffocation, anoxia)	Partially within Special zone for seabird protection
Kwinte swale	Monitoring of the sand/gravel ratio	Scouring (Lineations)	Abrasion (loss of substrate/complexity)	Habitat Directive Area
Goote Bank swale	Monitoring of the sand/gravel ratio	Erosion, Remobilisation, Scouring (Lineations), Smothering, Fine sediment entrainment, Overtopping	Abrasion (loss of substrate/complexity)	Habitat Directive Area
Thornton Bank swale	Monitoring of the sand/gravel ratio	Erosion, Remobilisation, Scouring (Lineations), Smothering, Fine sediment entrainment, Overtopping	Abrasion (loss of substrate/complexity),	na
Northern exploration zone	unknown	unknown	unknown	na
*Note	*Seascape* ref. to Pittman et al. (2011)	Interpreted with the same order as "Potentially influencing nearest anthropogenic activity" tab.	*Ref. for terminology - Miller et al. (2002 and references therein).	Ref.: Belgian MSP https://www.health.belgium.be/sites/default/files/uploads/fields/fpshealth_theme_file/19094275/Summary%20Marine%20Spatial%20Plan.pdf

Considering the importance of spatio-temporal backscatter data comparability, it was demonstrated how a standard practice of surveying a sufficiently stable reference area (regarding its well-documented backscatter signature - Roche et al., 2018; Fig. 6.1) at the start of oceanographic campaigns and by maintaining rigorous standards in acquisition and processing (Chapters 3, 4 and 5), the value of the backscatter measurements in respect to future change detection studies is greatly enhanced, making seamless backscatter coverage mapping possible when multi-source datasets are used. The reference area allows controlling the backscatter measurements repeatability. Standardisation and repeatability are especially important in view of the variety of sonar manufacturers, backscatter processing platforms, and the lack of pre-calibrated multibeam systems. Because of these reasons, working with multi-source MBES backscatter datasets is a globally recognised challenge in the field of ASC (Hughes-Clarke et al., 2008; Lurton and Lamarche et al., 2015; Lacharité et al., 2018; Misiuk et al., 2018; Weber et al., 2018).

The advantages of relying on a stable natural reference area are manifold. In this thesis, it allowed controlling the consistency of the backscatter measurements acquired by the system installed on the main vessel operating the monitoring programme, thus allowing the spatial and temporal comparability, and it allowed comparing measurements from different systems and platforms, producing a harmonised backscatter dataset that was calibrated against the ground-truth data through automated image-analysis approaches. This approach aligns with recent backscatter mapping literature where it is demonstrated that the predictive accuracy of classification models applied to harmonised backscatter datasets (as in Misiuk et al., 2018 and Chapter 3), as oppositely to analysing the datasets in isolation and combining the results post-hoc (as in Lacharite et al., 2018), can be significantly enhanced, as well as resulting in seamless datasets without noticeable “edge effects” affecting image-analysis. Importantly, control of the repeatability is inherently linked to subsequent change detection applications as measurements become also directly comparable in the fourth dimension: time (Chapter 4 and 5).

Systematic data acquisition over a reference area at the start of oceanographic campaigns contributes to meeting the objective of setting up a successful seabed substrate monitoring strategy in Belgian waters, as well as promoting the value of such an approach between neighbouring countries and farther afield. As specified in Chapter 1a, the value of the backscatter datasets used in this thesis (and in general) could be further enhanced by performing a full (absolute) cross-calibration with data acquired from a fully calibrated system (as in Eleftherakis et al., 2018 – in Belgium, this is an ongoing project in partnership with IFREMER and the CSS and NOC, whose ambition is to build a network of at-sea stable and monitored reference areas – see Roche et al. (2018) for details).

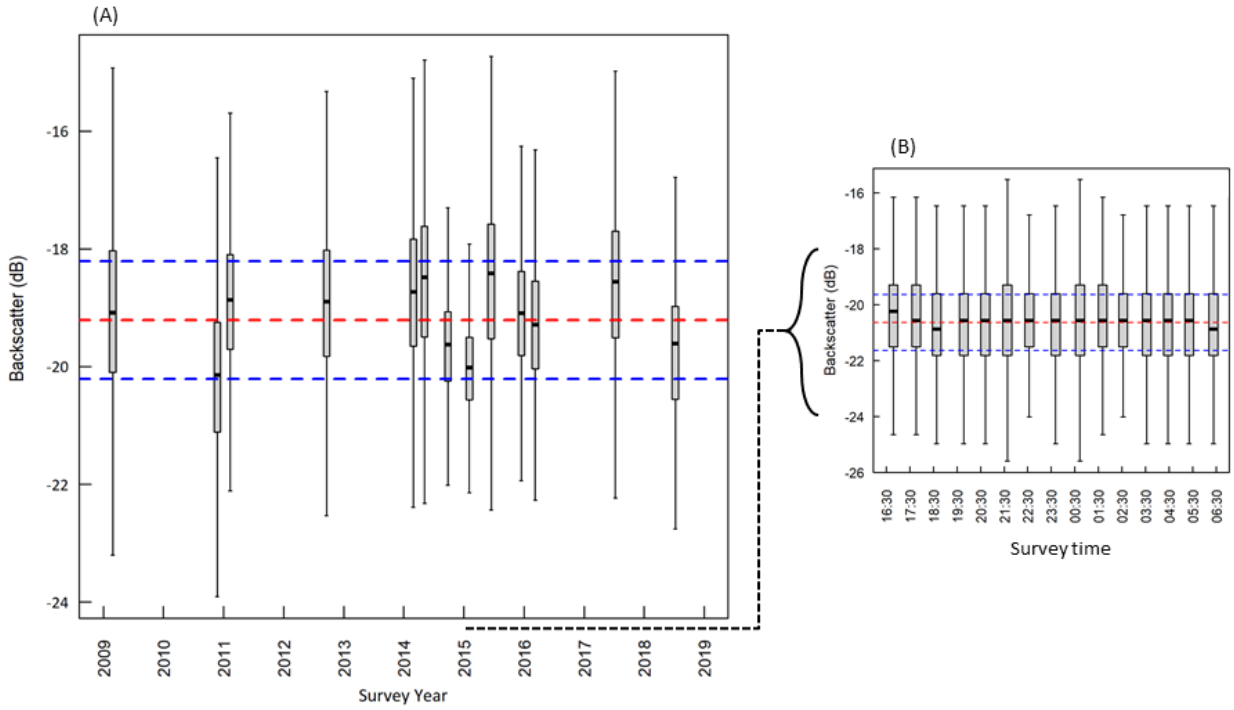


Figure 6.1 – Long and short-term stability of the backscatter natural reference area for the BPNS (Kwinte). Surveys acquired by the EM3002D system hull-mounted on RV *Belgica*. along predefined navigation lines and using strictly standardised acquisition and processing parameters. A) Long-term stability (yearly surveys). B) Short-term stability (13h tidal survey - Ref. Ch. 4). Boxplots: boxes denote the data contained within the 1st and 3rd quantiles and the whiskers the full non-outlier range of the data. Red line is the overall timeseries mean backscatter, whereas blue lines are the ± 1 dB Kongsberg manufacturer transducer sensitivity. This figure is important to the measurements used in this doctoral thesis and it shows how under standard acquisition and processing workflows, the data stability and repeatability can be assessed, fully justifying the applications presented in Chapter 3 and 5, based on the assumption of temporal comparability. Data in A (until 2016) and B: courtesy of Dr. Marc Roche, Continental Shelf Service, Federal Public Service, Belgium.

6.1.1 Data integration

Central to advancing the discipline of acoustic seafloor classification (both for mapping and monitoring) is the move towards automated classification methodologies deriving reliable (i.e. spatially-explicit, repeatable and accurate) maps showing the location, type and extent of benthic substrates (Anderson et al., 2008). In this regard, Chapter 3 fulfilled further the objectives set out under the first step by providing an implementation of two statistically-driven approaches establishing relationships between benthic substrate observations (from an array of ground-truthing sampling gears) and the suite of MBES predictor variables (backscatter, bathymetry and derivatives). Here, as in Chapter 5, supervised and unsupervised routines were compared, demonstrating the potential of automated approaches as a significant way forward compensating the lack of objectivity and repeatability of manually derived/digitised maps (e.g. Kaskela et al., 2019), that are often limited to confidence estimates (Butman et al. 1992; Verfaillie, 2008, Ch. 7.2, p. 150 and

references therein), hence lacking quantification of spatial uncertainty (accuracy) and detail.

There exists a paucity of research dedicated to the comparison of classification routines despite the importance of identifying which ones from the suite of available classifiers that can produce most accurate results (see Chapter 2 - *Recent investigations on seafloor mapping using automated image-analysis*). The choice of classifier remains a subject of serious discussion (Anderson et al. 2008; Stephens and Diesing, 2014; Calvert et al. 2015; Snellen et al., 2018). Montereale-Gavazzi et al. (2016) for example, compared an array of classifiers on a same dataset of very-high-resolution backscatter data (0.2 m) acquired in a very-shallow (5 m) tidal channel in the Lagoon of Venice (Italy). Amongst others, the study found that a relatively simple clustering-based procedure (i.e. involving the direct clustering of the backscatter gridded data only) superseded rather complex machine learning approaches in terms of thematic detail and accuracy: this motivated the comparison between k-means (referable to as a former-generation clustering technique) and Random Forest (referable to as a state-of-the-art machine learning classifier) throughout Chapters 3 and 5 of this thesis. The comparison revealed high similarity in performance in Chapter 5 whereas considerable differences were found in Chapter 3 (in terms of accuracy metrics). This implies that comparative studies claiming a “best” algorithm over others should be interpreted cautiously as it becomes clear that the performance of certain classifiers varies with study area and seafloor type under investigation, with characteristic underlying data structures. This encourages the scientific community to consider the comparison of classifiers as a standard prerequisite of any mapping investigation. While still at the nascence of automated classification applications, the choice of the classifier in a study is justified when several studies (comparative and not), from different geographical settings and seafloor compositions report on the accuracy of one particular classifier. In this respect, and accordingly to the most recent research (Stephens and Diesing, 2014; Li et al., 2016; Herkül et al., 2017; Gazis et al., 2018; Ierodiaconou et al., 2018; Porskamp et al., 2018; Turner et al., 2018), the Random Forest classification approach was found to be well suited for the integration of multibeam acoustics data, enabling mining of complex multivariate data structures, identifying relevant features and producing cross-validated uncertainty estimates.

It is important to note here that, while both classification approaches were able to resolve roughly similar broad-scale features, the two approaches resulted in considerably different maps and accuracies. These differences are rooted in the fundamental differences of the input data used, and by the different approaches themselves. The clustering approach was based on backscatter data only, whereas the influence of adding bathymetry and derivatives was tested only with the supervised approach. Indeed, the clustering approach can be improved upon inclusion of multiple input predictor variables in combination to PCA as demonstrated by many (Preston et al., 2001; Gavrilov et al., 2005; Eleftherakis et al., 2012).

However, this was not the purpose in this investigation as the interest was in appreciating whether a fast implementation of a widely used unsupervised clustering method, could produce satisfactory results based on backscatter alone as well as exploiting class-finding wrapper functions of the method.

Recent research by Snellen et al. (2018) and co-workers, showed how k-means clustering applied to backscatter alone had the tendency to yield rather evenly sized clusters, justifying the good performance in Chapter 5 where the study area had roughly evenly distributed classes and the poorer performance in Chapter 3 where the extent and structure of the data was considerably different. By comparing a range of accuracy metrics for a range of predictive models based on (1) backscatter alone, (2) backscatter and bathymetry and (3) backscatter, bathymetry and further relevant predictor variables, the influence of MBES-derived variables on supervised model predictive accuracy was tested. Key findings support the notion that bathymetry and backscatter and the set of predictor variables identified as relevant, are valuable surrogates for benthic substrates (Wilson et al., 2007; consequently, linking to surrogates of benthic biota – McArthur et al., 2010) and enhance the performance of acoustic classification (Eleftherakis et al., 2012). For instance, bathymetry-roughness, slope and backscatter Moran autocorrelation, were crucial in discerning between sandy Mud (in the nearshore surveyed areas) from hypothesised water-saturated Sand (on top of offshore sandbanks): together they allowed identifying areas of similar backscatter signatures, though morphologically distinct and differently organised in terms of textural variability. These predictor layers inevitably enhance seafloor characterisation by applying specifically to sedimentary processes, dictating the susceptibility to depositional and erosive processes (i.e. slope) and describing the textural complexity of the seafloor (i.e. Roughness – rough/rocky, flat/smooth - Wilson et al., 2007), compensating the lack of discriminatory power of backscatter alone.

While a multi-scale approach was not tested in this research, it has been demonstrated that the importance of predictor variables and the overall classification performance could be enhanced by including multiple resolutions (e.g. Wilson et al., 2007; Misiuk et al., 2018; Porskamp et al., 2018). Subsequently, the importance of the predictor variables found in this research must be interpreted with respect to the scale with which they were derived (3 x 3 neighbour on 5 m resolution pixels in Chapter 3 and the same in Chapter 5 though using 1 m grids). Figure 6.2 shows an example of the effect of resampling grid resolution on thematic accuracy, showing the effect of upscaling the grid in respect to the scale acquired by the sampling effort.

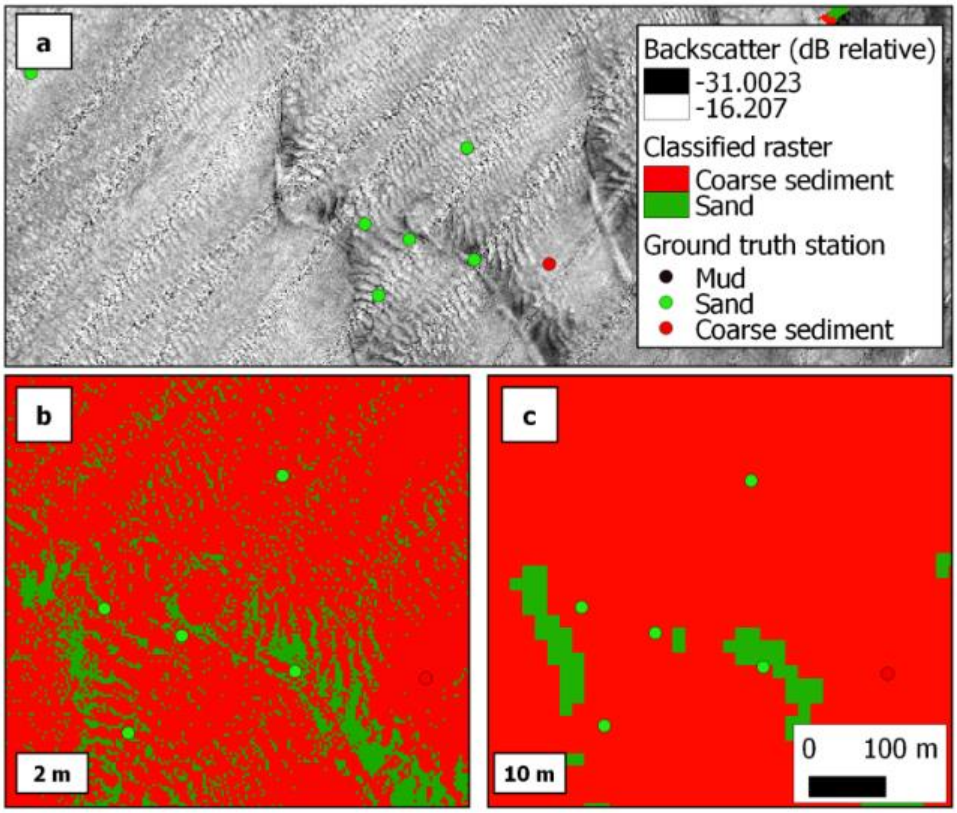


Figure 6.2- – Illustration of the dependency on scale of the accuracy metrics. A) Detail of a 2 m backscatter mosaic from a part of the ST1723 survey in the Hinder Bank area. B) Classified area into coarse sediments (red) and sand (green) at 2 m horizontal resolution and C) same classification with a 10 m horizontal resolution. Averaging the grid results in loss of fine-scale information that matches the scale of the sample observations, having implications on model training and thematic accuracy.

The choice of a 5 m minimal mapping unit in the investigation, as presented in Chapter 3, was based on considerations of the increasing size of the MBES footprint toward the outer portion of the swath and the survey lines' overlap. In 50 m of water (roughly comparing to the depth of the surveyed northwesternmost part of the BPNS) and considering the RV *Belgica* EM3002D system with 1.5° beams, the "pixel-size" from the nadir to 75° increases by a factor of three ranging from 1.5 to 5 m (Kenny et al. 2003). A finer gridding strategy would have led to the creation of artificial data in the mosaicking phase, possibly leading to excessive ambiguity in the interpolation (Gavrilov et al., 2005). In this regard, though at the cost of increasing surveying times, increasing the survey lines overlap would allow deriving higher-resolution datasets (especially in the deepest offshore areas), subsequently allowing analysing finer scale processes (i.e. topographic roughness), and predicting more detailed substrate classes (Misiuk et al., 2018; Porskamp et al., 2018). For example, a better discrimination of classes/features such as "outcropping gravel" and "biogenic structures" (e.g. *Sabellaria* sp. reefs), known to have characteristic fine-scale (bathymetric)topographic roughness (Jenkins et al., 2018). Besides this, a 5 m mapping unit satisfies imaging of both site-specific and regional patterns.

6.1.2 How many classes? With which classification scheme? "What can my sonar see"?

Investigation of the fitness of classification schemes and the class separation potential in respect to a given classification scheme (ultimately dictated by the sediment-discriminative ability of backscatter data at the frequency in use) are key aspects of ASC (Anderson et al., 2018; Strong et al., 2018). An important aspect of the data-integration phase covered in Chapter 3 related to investigating the sediment-acoustic relationships and, in light of the available ground-truth data, identifying the potential physical support for the relationships observed, and the implications these may have in view of prescribing (fitting) substrate classification schemes, shedding light over important classification limitations. Indeed, the "hydro-acoustic map-maker" has to bear in mind that the patterns used for classification relate to "acoustic diversity" which can often be related to, but not uniquely reflecting, the "real-world" physical character of the seafloor (Anderson et al., 2007). The Folk (1954) ternary classification (based on the relative proportions of mud, sand and gravel) was the target of interest here as this classification scheme is broadly applied at the European level: both by various international seafloor mapping initiatives (Kaskela et al., 2019, and references therein) and by other site-specific mapping investigations (e.g. Diesing et al., 2014; Gaida et al., 2018, Fogarin et al., 2019), together promoting the transboundary harmonisation of thematic mapping products in European waters.

The following important limitations were identified:

1) Identifying the number of classes in backscatter data via statistical-clustering aids is challenging due to the inherent noisy nature of backscatter and due to the nature of the clustering method itself.

2) Prescribing a classification scheme is equally challenging as it may not entirely match the backscatter discriminative ability of the sensor and its operating frequency.

Evaluating class separability in respect to a given classification scheme and gaining insights into the sediment-discriminative ability are interrelated issues, and their interpretation is not a trivial task. At first, it is worth noting that the inherently noisy nature of backscatter data (Jackson and Richardson, 2007) makes this aspect as challenging as intriguing in the context of a classification problem. As reported by Snellen et al. (2018) "*the natural fluctuation of backscatter can superimpose the backscatter variation due to different seabed properties*". Indeed, the stochastic nature of the backscattering phenomenon leads to random fluctuations of the Echo level (Malik et al., 2018). By logic, attempts at estimating the number of classes in a backscatter dataset is best approached unsupervisedly (e.g. sum of squared distances, silhouette coefficient – Chapter 3 and 5) where no *a priori* knowledge of the inherent data-structure exists, and the natural groupings of the data are sought

after by the method. By this approach, the investigator relies on the efficiency of the unsupervised algorithm at mining the "real" number of classes, or rather, those that the backscatter can support, performing a final semi-automated match with h classes. On the contrary, a supervised classification approach informs the classification with an *a priori* number of classes, informed by the e.g.: sedimentological or biological (or both) interpretation of the ground-truth data. Both approaches force the fit of a predefined classification scheme, either *a priori* (supervised) or by seeking to find the optimal match (unsupervised).

In practice however, the attempt to automatically (unsupervisedly) identify the number of classes in the backscatter data by applying statistical-clustering aids, the underestimation of the number of clusters can be expected given both the noisy/fluctuating nature of backscatter measurements and the way the clustering algorithm works (here referring to partitive clustering k-means based on Euclidean distance): in this case the minimisation of Euclidean distances of data points from their cluster centroids in a k-means analysis, inevitably leads to favour the identification of rather symmetric and compact patterns (in the feature space), which may not necessarily reflect the backscatter data structure (as also observed in Snellen et al. 2018). This hinders a clear identification of the number of clusters for most statistical aids built on non-overlapping data structures, formulated on the assumption of clear peak-separation in the data (i.e. the numerous indices presented in Milligan and Cooper (1985), often referred to in the seafloor mapping literature).

In Chapter 3, the statistical-clustering aids applied to identify the number of classes were poorly performant likely due to this reason and converged to (weak) optimal solutions of 3 to 4 clusters, making the objective selection of the *optima* somewhat challenging. Besides this, the exploratory data analyses using boxplots comforted the notion that both 3 and 5 class solutions could be reasonably discerned in the backscatter data (following class aggregation for which fundamental and potential explanatory reasons will be discussed shortly hereafter), and further enhanced by the inclusion of relevant explanatory morphometric variables (or predictor layers) in the supervised classification approach. Based on these observations, for point (1), it may be deduced that statistical clustering-aids are not robustly applicable unless clear peak-separation is observed in the underlying data structure. Consequently, similarly to testing multiple classifiers, investigators are encouraged to test different "k-finding" aids and establish which can describe the underlying data best.

To improve this aspect of the mapping investigation in Chapter 3, and argument further the observations on "finding the number of classes in respect to the discriminative ability of the backscatter data", another approach was tested. The fitting procedure proposed by Simons and Snellen (2009; 2.2.1) was followed and tested on the backscatter dataset of Chapter 3 (the overall seamless map referenced at oblique beam angles - $BS_{45°}$). The fitting procedure approaches the issue of finding the number of classes by a minimisation problem, testing against the reduced χ^2 the number of m Gaussian probability density functions (PDF) that can be fitted to

the backscatter data PDF. The optimum is reached when adding Gaussians does not improve the reduction of the χ^2 criterion any further. The analysis identified the goodness of fit at six classes (Fig. 6.3). The residuals show well defined structures suggesting that either more PDFs could be fitted (though the minimisation used fails at improving this situation; note that two of these PDFs are too small to be seen) or that the Gaussians do not represent the data-structure adequately. This further approach approximates to the class separation potential observed in Chapter 3 by both boxplot (empirical/expert differentiation) and clustering-aids (statistical differentiation).

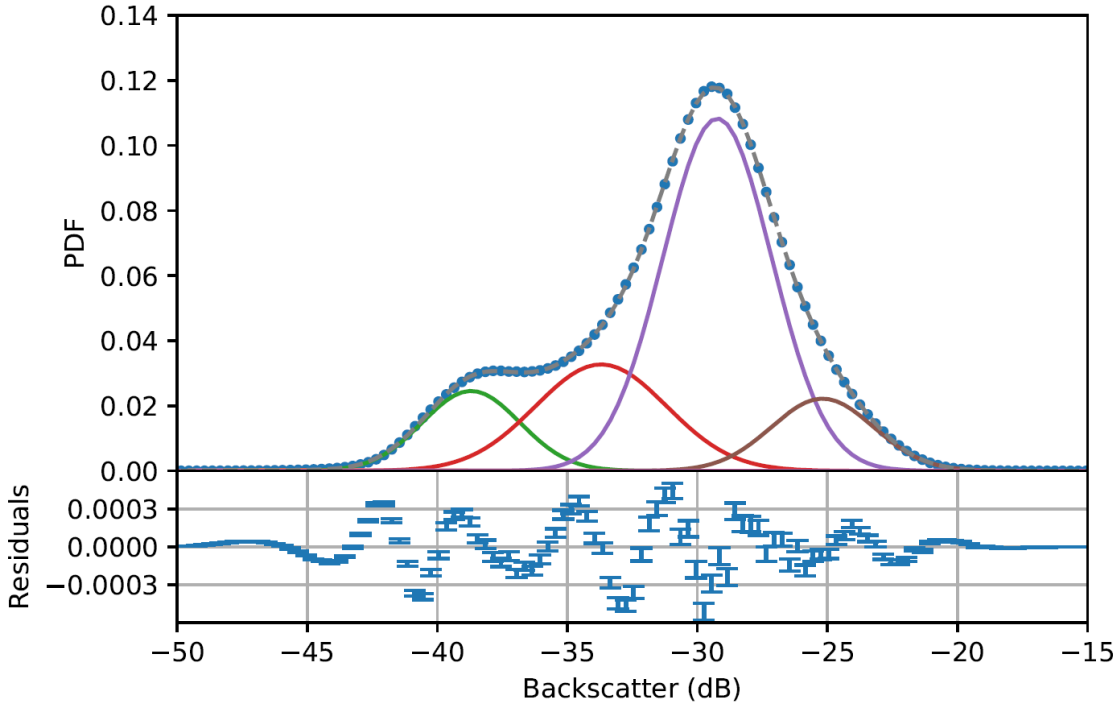


Figure 6.3 – Fitted PDF of backscatter data per bins of 0.3 dB (normalised histogram). Backscatter data from Chapter 3. The starting mean points for the fit where evenly distributed across the range of x. Standard deviation was set between 0 and 5 and the amplitude between 0 and 0.6. The search zone (prior/bounds) was set between -50 and -10 dB. The degrees of freedom are the number of parameters in the model (3 for each Gaussian). The reduced χ^2 is defined as the $\frac{\chi^2}{\mathrm{df}} - 1$.The error is given by the square root of the number of samples per bin. Mean, SD and amplitude for each class are (left to right): m_1[-44.9669, 2.4946, 0.0010], m_2[-38.6168, 1.917, 0.1218], m_3[-33.864, 2.499, 0.2115], m_4[-29.2255, 2.0510, 0.5482], m_5[-25.2876, 1.9943, 0.117], m_6[-20.0428, 2.5, 0.0001]. m_1 and m_6 are artificial. It is interesting to observe that the median (and mean) values of the Folk 5 boxplot in Chapter 3 fall within the acceptance regions of this analysis. The fitting procedure is as in Simons and Snellen (2009) except for the search range of the standard deviation that here was arbitrarily set. As such no rigorous physical prior is assumed in this modelling approach.

Aggregation of classes is a great issue in ASC and has been encountered by many investigators (e.g. Diesing et al. 2014; Buscombe et al. 2017; Gaida et al. 2018; Fogarin et al. 2019). Its implication is the production of maps which are relatively less informative (complex) then they could potentially be. For point (2), fundamental potential physical reasons exist hindering the sediment-discriminative ability of

backscatter (here acquired at 300 kHz) and therefore impact on the assignment of prescribed categories to the data. It is important to be aware of these issues throughout the map-making process.

Firstly: gravel content (%) exhibits a strong yet non-linear relationship with backscatter intensity (Goff et al., 2004; Diesing et al., 2014). Even a minor portion of gravel (~5 %) can considerably influence the acoustic response, leading to a masking effect of the finer sediment fraction, regardless of its nature (or sand:mud ratio considering the Folk classification); this phenomenon was observed in Chapter 3, leading to aggregation of mixed and coarse sediment classes to fit the prescribed classification scheme. This dispersion effect was particularly evident in the Ostend study site where spatially concurrent box-core and SPI samples greatly improved the interpretation of the acoustic response (ref. Fig. 3.24b).

Secondly: based on theoretical laboratory work by Ivakin and Sessarego (2007- based on well-sorted homogenous and degassed marine granular sediment), recent field measurements in similar sedimentary seafloor environments by Buscombe et al. (2017), Gaida et al. (2018) and Snellen et al. (2018) as well as from observations gleaned from the exploratory data analysis of Chapter 3, unexpected acoustic responses could be expected when the ratio between the sediment grain mean diameter and the acoustical wavelength exceeds 1. For a 300 kHz sonar system, this equates to grains with a mean diameter of approximately 5 mm. Beyond this grainsize ($d/\lambda > 1$), the scattering is dominated by discrete scatterers, and the sediment grains can no longer be considered as a continuum ($d/\lambda < 1$): a transition from positive to negative correlation or a decrease in magnitude of strength in correlation with backscatter intensity can be expected in this scenario.

Thirdly: and most challenging to estimate from the ground-truth data, are non-anticipated phenomena such as the effect of volume scattering from discrete inclusions in the sediment matrix (Urgeles et al., 2002; Ivakin, 2008), sediment layering and gas inclusions (Williams et al. 2009) and bioturbation (Urgeles et al., 2002; Gorska et al., 2018). The latter are possibly the most prominent factor influencing high-frequency backscatter though remain mostly unaccounted by the lack of appropriate ground-truth data.

The strong relationship between D50 (for sediment grains up to ~ 550 µm and well-sorted samples scored as “clean” – i.e. predominantly catheterised by the sand gains only) and backscatter intensity observed in Chapter 3 provides empirical evidence that fine (≤ 250 µm), medium (≤ 500 µm) and coarse (≤ 2 mm) sands are benthic substrate categories (i.e. sensu Wentworth, 1922) that could be predicted and mapped based on backscatter data (Goff et al., 2000; Collier and Brown, 2005; De Falco et al., 2010). The same expectation can be drawn from the work of, for example, Jackson et al. (1986) whose composite-roughness backscatter model (part of the APL-UW, 1994) demonstrates the feasibility of acoustically predicting well-sorted, mostly unimodal and homogenous marine sands (see Chapter 2). It remains

true that this approach is limited to ideal and canonical configurations of the marine substrate (Lamarche et al., 2011; Lurton and Lamarche, 2015), rarely occurring in nature (Anderson et al., 2007).

Because complexity of marine sediments is the rule rather than the exception, several geotechnical and geoacoustical parameters are found to orchestrate the acoustic response (from an extensive literature review presented in Anderson et al. (2007; pp. 9), it is found that up to 80 parameters have been used in the attempt to "holistically" describe the physical structure of the water-sediment interface and relate these to the acoustic response). Fortunately, several of these parameters co-vary with one another and many investigations have demonstrated empirically the relation of backscatter mosaic data with fewer parameters readily accessible from the ground-truth data (e.g. Collier and Brown, 2005; Ferrini and Flood, 2006), rising the potential to interpret the high-frequency backscatter data for substrate characterisation in the operational environment, where the possibility to carry out detailed and controlled measurements required for a rigorous interpretation of the acoustic response, is often hindered by logistics inherent to *in situ* work.

Because of this, it is crucial to keep in mind the limitations and the possible erroneous interpretation this may cause when using backscatter only for classification with limited access to ground-truth data.

The striking similarity of the acoustic response from the uppermost portion of sandbanks (Hinder Bank area) and muddy (Ostend disposal site) areas observed throughout Chapter 3 (as well as by previous seafloor mapping studies for the same areas and based on similar interpretative techniques), is a good example of this. Despite the following interpretation remaining anecdotal, it was hypothesised that the shared acoustic similarities between these two morphosedimentary distinct features were due to the sand's high water-saturation and the high sediment transport[1] at the time of data acquisition on top of the sandbanks, leading to an acoustic response comparable (in terms of gridded dB values) to that of the unconsolidated, fluid-like muddy seafloor at the Ostend study site.

Assuming this interpretation is correct, this would demonstrate the limitations of considering only the sediment grainsize and/or only the relative proportions of sediment fractions when interpreting the acoustic signal, fully justifying the integration with morphometric variables to enhance the discriminative power of the classification and recognising the intricacy of "acoustic diversity", dependent on limitedly accessible geoacoustical and geotechnical parameters.

This is especially the case in the BPNS and for the high-frequency sonars used in shallow waters where the seafloor is characterised by a continuum of grain sizes and morphological features at various scales (i.e., sandbank tops, swales and offshore planes), and collectively featuring elements which are both below and above the acoustic wavelength. While the influence of large-scale (above-beam)

morphology can be corrected based on a bathymetric DTM (as it is the case for the data used in this work), the strong (sub-beam) topographic component in the backscatter signal at 300 kHz, possibly represents the most severe limitation on the achievable precision of an acoustic sediment classification. Ferrini and Flood (2006) provide an insightful multivariate approach to a backscatter study in which sub-beam topographic roughness was used as a variable. Besides finding that sub-beam topographic roughness is a key factor, at least for sandy, siliciclastic seafloors, the study concluded that *"Observation of characteristics that are important with respect to the acoustic properties of sediments and signal penetration in addition to grain size and roughness measurements may also be needed (e.g. bulk density, sound velocity, gas content, shell characteristics)".*

In light of these observations, finding that increasing the number of categorial classes decreases the predictive accuracy of the classification model and the subsequent need of recurring to class aggregation, comes as no surprise and while it may be reasonable to assume that the decreasing performance relates to "*an increase in the complexity of the classification task*" due to an increase of rules/decisions and a decrease in the number of samples per category (e.g. Porskamp et al. 2018 and as surmised in Chapter 3), it is crucial to identify the physical support that may, at least partly, explain why is so when no clear conclusions can be drawn.

[1]Indeed, surveys acquired on top of sandbanks are always acquired during high tide (otherwise being hazardous for navigation), therefore often coinciding with moments of strongest current velocity and sediment transport. As it will be discussed later (and ref. to Chapter 4), near-bed sediment transport remains a poorly quantified factor in respect to sonar performance and could be an explanatory factor supporting this observation in this study and in previous mapping and classification studies by Roche (unpublished), where coincident observations were made.

6.1.3 Ways forward

The notion that backscatter on its own, at least at a single frequency, may not suffice in describing the natural complexity that characterises the seafloor, is not surprising considering the observations raised up to here. Nevertheless, a plethora of exciting and innovative approaches, ranging from novel data-structures, improving design of sonar systems, to integration with machine-learning algorithms (i.e. Chapter 3 and 5), leaves us with numerous interesting possibilities that we can exploit to our favour in the challenging task of making sense of the submerged world.

Besides the recommendations set-out in Chapter 3, this brief section lists a set of recommendations that could be addressed in future investigations to improve the discrimination potential of backscatter data registered by multibeam bathymetric echosounders:

1) According to Ivakin and Sessarego (2007) ambiguities in field studies, particularly those in shallow and dynamic coastal environments, come as no surprise as the unpredictability of the environmental conditions challenges the reduction of uncertainty in the interpretation of acoustic data (cf. Chapter 4), even when time-consuming and labour-intensive measurements are put in place (e.g. the Sediment Acoustic Experiments [SAX04, 99] reported in Williams et al., 2009). **Controlled laboratory experiments**, as those carried out at the French Laboratory for Mechanics and Acoustics tank-facility (http://www.lma.cnrs-mrs.fr/ - e.g. Korakas et al., 2008), addressing the study of the acoustic response from replicated naturally occurring conditions, would prove invaluable (and considerably less expensive) in view of the interpretation of data acquired in the field. Alternatively, this exacerbates the need for system's absolute calibration, reducing the uncertainty associated with the comparison of laboratory and field measurements.

2) The **specific design of multiparametric ground-truthing gears**, combining optical, physical and geotechnical observations of the area being sampled (including acoustic signalling devices to improve the positioning of the sample), would inevitably improve the parametrisation of the samples acquired, improve cost-efficiency of sampling efforts and enhance the characterisation of backscatter data in respect to the naturally occurring complexity. As mentioned in the Discussion section of Chapter 3, while the remote sensing technology has drastically evolved over the past two decades, along with processing capabilities of modern computers and the improvement of sonar-image processing (e.g. compensating angular dependence and dealing with the strong nadir banding effects), except for the advancing underwater videographic technology, rapid groundtruthing approaches have remained relatively underdeveloped, with much room for improvement.

3) Moving **from hard/crisp** (i.e. discrete/categorical) **to fuzzy** (i.e. numerical/continuous) **classification** of the remotely sensed data, i.e. deriving maps independently of classification schemes, could overcome the issues inherent to fitting a prescribed scheme (Strong et al., 2018), as well as improving our understanding of environmental gradients (ecotonal patterns). This may prove particularly valuable considering that benthic species may not necessarily "neatly" fit into a sediment category (Mitchell et al., 2019). Recent seafloor mapping studies demonstrated the potential of using the Random Forest classifier for regression (continuous data), as oppositely to classification (categorical data). Misiuk et al. (2018) for example, predicted mud, sand and gravel percentages over a large-scale multi-source backscatter dataset using a tree-based classifier. In turn, the predictions were recombined to represent categorical classes. Gazis et al. (2018) used Random Forest (regression) to predict the percent cover of manganese nodules, further demonstrating the potential of such an application.

4) Improvements of the discriminative ability of the backscatter acquired by the system in use, could come in from an approach referred to as Hyper-Angular-Cube (HAC) initially proposed by Hughes-Clarke (1994), implemented by Parnum (2007) and proposed again in a recent study by Alevizos and Greinert (2018). This method refers to a "stack" of **multi-angular backscatter grids**. Generally, normalisation of the angular dependence during backscatter mosaic production entails a normalisation of the data to an angle or a range of angles. Recall from Chapter 2 (and 3) that this is generally applied to the set of angles where the sediment class separability is maximum, i.e. around the "plateau" angular region, nominally at 45° (effectively averaged between 43° and 47° - or between 30° and 60° in Fledermaus Geocoder engine – Chapter 4). The HAC then proposes to produce several backscatter mosaics, each normalized at a different incidence angle. This has tremendous advantages in view of solving a classification problem as it enhances the resolution of the angular response backscatter by gridding the angular data as a function of the bathymetric resolution (otherwise limited to approximate the size of the window used for its computation – i.e. a set of consecutive pings covering a large portion of the seafloor at port and starboard sides of the swath), keeping at the same time the angular information and the fine-scale bathymetric resolution. Different sediments are expected to result in different dB values across the angles of incidence (and as exemplified in Chapter 3 – Modelling the angular response).

5) Recently, over the past three years, MBES technological development has seen the appearance of **multi-frequency systems**, able to shoot simultaneously 100, 200 and 400 kHz sound waves (e.g. for an R2Sonic 2026 MBES). Use of these sensors for benthic substrate characterisation is in experimental phase and the potential for increased discriminative ability has

been reported in a number of recent publications (e.g. Feldens et al., 2018; Gaida et al., 2018). The logic behind a multi-frequency sonar system, homologous to a terrestrial multi-spectral, multi-band system (e.g. Shaw and Burke, 2003), is the enhancement of the "sonar's seafloor perception" in respect to the use of multiple-sized wavelengths. By allowing multiple "perception modes", i.e. different degrees of acoustic penetration into the sediment and sensitivity to roughness of different scales, the acoustic signature for a given seafloor type can be richer in information, possibly enhancing class discrimination. Multi-frequency systems may be particularly interesting to improve the discrimination of seafloors with gravel since varying frequencies (far enough between each other) would allow simultaneously perceiving the seafloor as a continuum and as individual/discrete scattering features (i.e. pebbles/cobbles/boulder). It must be noted however, that the ground-truthing effort would increase here, as the subsurface would require a more rigorous characterisation compared to samples targeting the characterisation of the immediate seafloor, expected to primarily influence high-frequency (i.e. 300 kHz) backscatter data.

6) A final, perhaps crucial observation, relating to the very first step of the backscatter data processing chain and the conversion from "raw" amplitude to backscatter strength (occurring in the built-in backscatter acquisition mode of a given echosounder), follows (after Fonseca et al., 2019): What is referred to as backscatter, when working with MBES systems, relates to an ensemble average of acoustic backscatter signals. Averaging over n backscatter samples is necessary to reduce the random uncertainty of the backscatter signal, related to its inherent statistical fluctuation (Malik et al., 2018). The backscatter strength per seafloor unit area (expressed in decibels and proportional to the square of raw amplitudes) is registered by a MBES under the form of "snippets" time series (describing the in-phase temporal dimension of the insonified footprint) or under the form of a single value per beam (e.g. depending on the system used derived from an average, maximum or else value of the sample's PDF). Averaging a set of backscatter samples thus implies averaging of squared amplitudes x^2. However, the traditionally used mean value may not be the most adequate statistic to summarise the PDF of the set of BS amplitude samples measured in the field by a MBES (generally following a Rayleigh distribution). Recently, it has been proposed to use the median as an alternative (and more robust) BS calculation method (Fonseca et al., 2019). The advantages of using this statistic over the mean are manifold. Importantly, compared to the mean, the median is a statistic that is less "inflated" by the number of scatterers that "contaminate" the footprint and hence more stable in respect to the number of raw amplitude samples used to derive the statistic. This opens a new and exciting perspective on approaches to backscatter data interpretation. For example, a sonar footprint incident on a predominantly sandy water-sediment interface and populated by few strong

scatterers (e.g. shell detritus), would considerably differ in terms of mean or median value. Clearly, a median value would better reflect the dominant substrate (i.e. surrounding the strong scatterers) whereas the mean value would be strongly influenced by the strong amplitude returns caused by those few scatterers and departing from the mainstream behaviour of the predominant surrounding substrate type. This would thus pose the advantage of accessing more "balanced" backscatter quantities, whose combined interrogation (in a classification problem) could considerably enhance the sediment characterisation. Alternatively, accessing other statistical measures beyond the mean and the median (e.g. min., max., mode, skewness and kurtosis) directly from the raw amplitudes, may lead to highly relevant backscatter variables, more than those statistical derivatives obtained from the gridded (CBI) backscatter data (Chapter 3). Here, the inclusion of such a statistical approach to the derivation of backscatter data to a multi frequency system, would inevitably increase the quantity of available information to the investigator since the different PDFs and statistics will differ between frequencies used: strong scatterers will "appear" at a higher frequency and "ignored" at a lower one. Separated-enough frequencies, correspond to different physical processes in the way the returning echo is constructed (how it interacts with the seafloor; its characteristic roughness, hardness and bulk density).

7) Finally, with regard to enhancing the overall classification performance, and that of specific classes, very recent studies are suggesting further exciting new possibilities proposing innovative variables derived from uncommonly used file formats (e.g. point clouds used to map Submerged Underwater Vegetation – SAV in Held and Schneider von Delmling, 2019) and previously unexplored backscatter textures and statistics (such as the Weyl transform backscatter indexes currently being studied on the basis of the Oostende backscatter dataset [Fig. 3.1b] in Zhao et al., 2019).

From an acoustic seafloor classification standpoint, responding to the question raised in Verfaillie (2008 - how good is the ground-truthing?) the goodness of the ground-truthing data can be appraised by how well the limited set of parameters allows a sufficient number of acoustic classes to be discerned and classified (both empirically and statistically). Given the difficulties of acquiring such information, especially when dealing with regional mapping efforts, targeting a variety of seafloor types, one has to bear in mind the combination of limitations and ambiguities that may arise. Consequently, the predictive models herein presented must be interpreted in respect to the degree of generalisation that was applied to their derivation (thematic categories). Nevertheless, although limitations do exist, they can be identified, and their effects appraised, and strategies proposed that allow current seafloor classification capabilities to be an invaluable asset in improving our understanding of the seafloor at relevant scales, both spatially and temporally.

The next section provides a brief summary of the ecological value of the maps produced and reiterates on the importance of spatial error of predictive models, introducing a model uncertainty visualisation approach of the most accurate map derived in Chapter 3. In turn, the section "*Seafloor monitoring using MBES: variability and change detection*" discusses further the ACD research of Chapter 4 and 5 in view of seafloor monitoring and identifies strengths and weaknesses, addressing future research, challenges and opportunities.

6.2 Ecological value of fine-scale predictive substrate models: on surrogacy

While the target of classification was on benthic substrate (abiotic) rather than benthic habitat *sensu stricto*, it was however felt that the following observations are of interest for a range of ecologically-minded applications and enhance the value of novel approaches to seafloor seascape and benthic habitat characterisation, to which the methodologies herein tested contribute, as well as having the potential to evolve into better benthic habitat maps assuming an improved coupling of geo-bio information.

From a benthic and marine ecology viewpoint, fine-scale predictive (thematic) models of the benthic substrate distribution are of high value as they provide ecological information at operationally relevant spatial scales, identifying seascape patterns, down the unit level of patches and corridors (Zajac, 1999; Pittman et al., 2011). The grain of the information we are reaching, is such that it can form the basis of an array of applications, of which importantly the identification and characterisation of benthic habitats (e.g. Todd et al., 2000; Kostylev et al., 2001; Diaz et al., 2004; Ierodiaconou et al., 2007; Brown et al., 2011; Montereale Gavazzi et al., 2016), the designation of priority areas for conservation (e.g. Ward et al., 1999; Ierodiaconou et al., 2007), prediction of species/assemblages distribution for conservation decisions (Guisan et al., 2013), various applications of ecological modelling (e.g. Ecological Niche Factor Analysis, Identifying fish habitat; Galparsoro et al., 2009; Guinan et al., 2009; Iampietro et al., 2005; Monk et al., 2010; Valle et al., 2011), change detection (e.g. Rattray et al., 2013; van Rein et al., 2011; Montereale Gavazzi et al., 2017; 2019) and applying novel concepts of seascape ecology (i.e. taking from the terrestrial landscape ecology: Boström et al., 2011; Pittman et al., 2011, 2007; Pittman and Olds, 2015; Wedding et al., 2011).

The notion that backscatter, bathymetry and their derivatives, act as surrogates that determine habitat availability and suitability, for example for epibenthic hard substrate communities, is supported by the maps and ground-truth evidence presented throughout this thesis. In terms of surrogacy, the presence of coarse sediments, ranging from shingle to coarse shell detritus and gravel, provides the structural complexity (i.e. hard substratum, crevices and roughness) needed for benthopelagic coupling and settlement, promoting rich and diverse epibenthic communities (McArthur et al. 2010). Figures 6.4 and 6.5 further corroborate and

support this observation showing a set of samples coincident within predominantly coarse substrate classes of the models produced in Chapters 3 and 5.

Preliminary testing of videographic sampling gears (purposely modified in this thesis to include laser pointers, metric reference scales, improved illumination and high-definition camera system) on RV *Belgica* and *Simon Stevin,* allowed imaging of abundant epibenthic growth, particularly in the Northwesternmost offshore plane, where soft coral *Alcyonium digitatum* was frequently observed (up to 10 ind. X m^2 frame). Using Van Veen grab sampling, important coarse shell detritus ground was identified in the Hinder Bank area, denoted by a speckled fine-scale pattern of sG and gS classes, providing habitat for important fish species the sand eel (*Ammodytes* sp.); a species whose habitat is threatened by marine aggregate extraction practices (De Backer et al. 2014). Areas of colonisation by *Lanice conchilega* (the sand mason polychaete worm) were also identified, though not dense enough to be acoustically imaged as in previous studies (e.g. Degraer et al., 2008; Van Lancker et al., 2012). Both Hamon grab and videographic sampling allowed identifying important bio-engineering species such as bio-encrusting polychaete aggregations (indet.) in the northern exploration area. Together, the potential for benthic taxa description provides a framework for follow-up benthic habitat mapping and fine-scale niche and habitat suitability modelling studies in the BPNS. Indeed, the areas of gravel mapped throughout this study, are spatially coincident with the "Potentially Ecologically Valuable" seafloor areas identified amongst others, in Verfaillie (2008; cf. pp. 156-161). In particular, the northern exploration area may represent an important area, resilient to anthropogenic pressures and ecologically valuable, deserving the consideration for new proposals of seafloor conservation and follow-up research campaigns.

In agreement with recent applications of underwater videography (Michaelis et al., 2019; Van Der Reijden et al., 2019), use of video frames was invaluable in hard-substrate seafloor areas where conventional gears are prone to failure, avoiding direct physical impact (Beisiegel et al., 2017; Chimienti et al., 2018) and increasing the sampling effort in space (i.e. from single-point locations to reconnaissance transects). As protocols for image analysis are improved, requiring time and biological expertise for proper identification down to the lowest possible taxonomic level and improvement of automated routines for sample description (e.g. https://dbuscombe-usgs.github.io/DGS_Project/), our ability to understand patterns of α and β diversity in relation to the seascape patterns observed in the acoustic imagery, will ultimately dictate our ability to predict γ diversity and improve the rigour of ecological descriptions of seascapes based on acoustic classification (Pittman et al., 2011; Wedding et al., 2011; Rocchini et al., 2016; Lacharite and Brown, 2019).

Recent studies in the Hinder Bank region have confirmed the persistence (resilience) of dense aggregations of epibenthic, flourishing and arborescent hard-substratum communities (e.g. Fig. 6.5) identified in pioneering studies of the same area (Gilson, 1907; Houziaux et al., 2008, 2011). Clearly, mapping of substrate type

in respect to the fine-scale geomorphology allows identifying correspondence with specific biological assemblages (van Dijk et al., 2012; Van Lancker et al., 2012; Van Lancker et al., 2017), highlighting the importance of spatial information of substrate and morphology derived from multibeam systems as key surrogates for benthic life. As such, accurate mapping of substrate and benthic features is critical to build the knowledge necessary to better inform the management of areas for environmental conservation as well as enabling investigators addressing in more detail answers about marine ecology and biology beyond safely and cost-effectively repeatable diving operations and over multiple spatial scales.

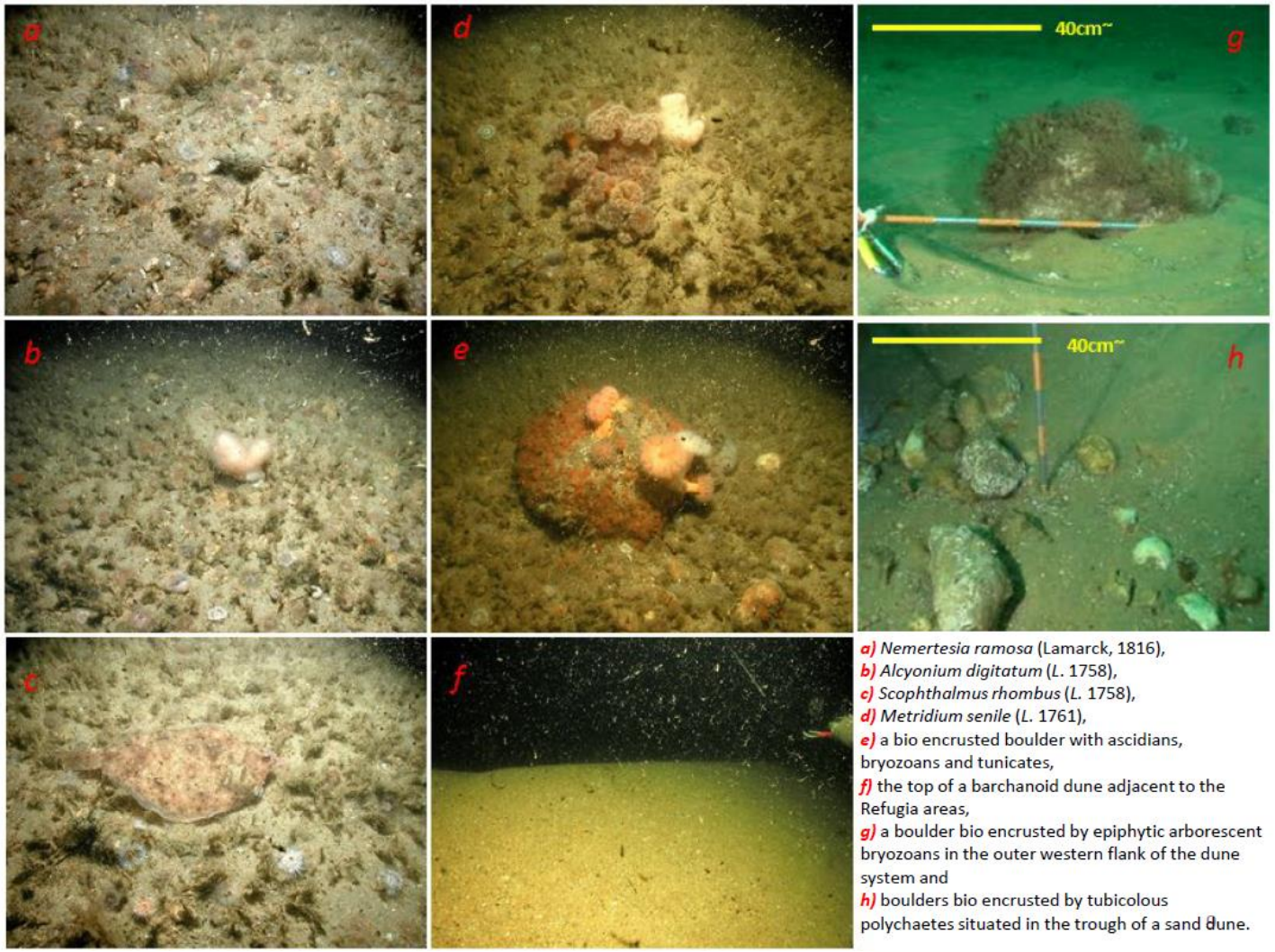


Figure 6.4 – Diver pictures in the trough of a barchanoid dune in the Hinder Banks (Chapter 3 and 5, Refugia zones). Courtesy of Alain Norro, Scientific Diver at Royal Belgian Institute of Natural Sciences.

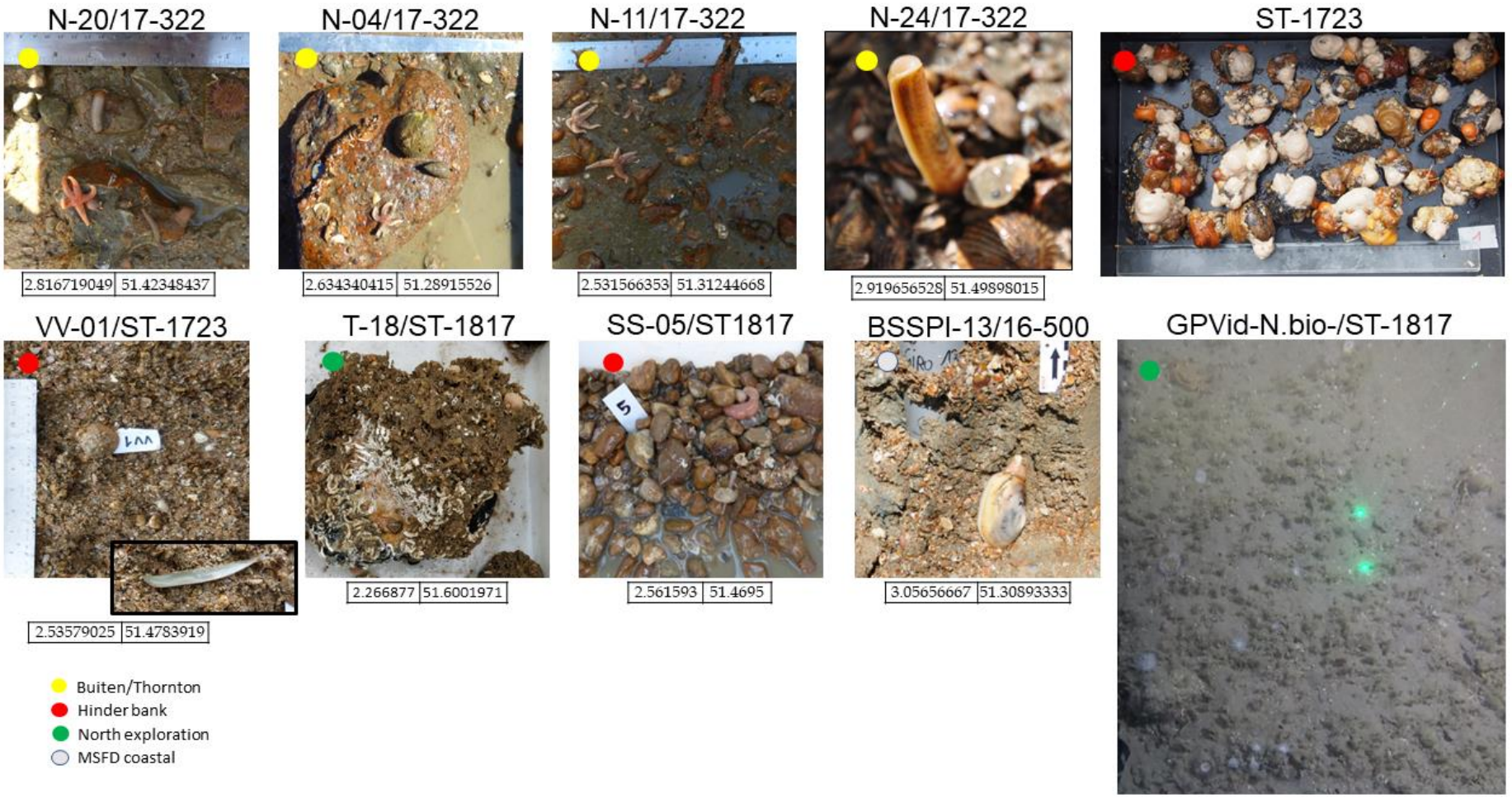


Figure 6.5 – Selection of samples indicative of strong associations between coarse/gravel substrate and macrobenthos. ID, Campaign code and Lat. Long. Coordinates are provided.

6.3 Quantification of spatial uncertainty

The kind of information conveyed by acoustic classification is particularly relevant when the quantification of spatial uncertainty is carried out and transparently communicated. Reporting accuracy is important not only for the map-maker, identifying where the model over and/or underestimated the predictions, but also for subsequent uses of the map by third parties. If these maps are envisaged to provide information to guide follow-up benthic habitat mapping studies, with particular reference to the poorly explored hard-substrate epibenthic communities of the BPNS (Houziaux et al. 2011, Verfaillie, 2008), then accuracy becomes a criterion whose reporting satisfies both scientific rigour and decision-making. The four questions raised in the *Thematic model's evaluation* in Chapter 3, describing the protocol of error estimation, have been addressed through a combination of a thorough matrix-based approach, visual interpretation and identification of limitations previously discussed. Crucially, reporting of error advances the scientific progress, in any discipline.

(1) What is the *error frequency*: how often does the map not agree with thematic *reality*?

This question was replied to by deriving various accuracy metrics (for a definition refer to Chapters 2 and 3). The commonly reported Global Accuracy is a useful "first-glance" metric to appreciate the goodness of the thematic prediction. For example, the 95 % confidence interval provides useful information: if the same sampling and classification scheme would be repeated for a large number of times, it is expected that in 95 % of the runs, the observed accuracy would be somewhere between those intervals with a 5 % risk that the true prediction would be beyond these intervals. Apart from the often solely quoted overall Accuracy (Story and Congalton, 1986; Congalton, 1991; and still to date), Chapter 3 estimated spatial errors using by-class accuracies, Chance Agreement k and the No Information Rate, together providing a solid error estimation protocol.

(2) What is the *nature* of the errors: which classes are not mapped correctly, and with which other classes are they confused with?

This question was replied to by discussing the discriminative acoustic resolution of the sensor (*sensu* backscatter), identifying possible explanatory factors of acoustic class dispersion.

(3) What is the *magnitude* of errors: how serious are they for a decision maker?

Answering this question is the inclusion of the No Information Rate to the protocol of error estimation. For example, considering the $RF_{FOLK++(5)}$ model produced, with an Accuracy of 74 % and a k statistic of 76 % one can appraise that the model has generally a high accuracy and most of the accuracy did not occur by

chance. Furthermore, the NIR of 37 % is the accuracy achievable if the model would always predict the largest class percentage in the data. A significantly low NIR confirms that the previous accuracy metrics are not inflated by the majority class. Furthermore, by-class accuracies (i.e. User and Producer) are especially important as they enable the identification of the reliability of the predictions for a given class. As an example, if a marine ecologist plans a sampling campaign based on the $RF_{++FOLK\ (5)}$ model, with an interest in the sM class, the map producer could claim a 50 % (Producer) accuracy of the map, but the user (the marine ecologist) would know that when visiting the field, only 33 % of the area predicted as that class can be reliably related to sM. Due to this, these maps can accurately inform a range of applications, including follow-up sampling campaigns.

(4) What are the *sources* of errors: why did they occur?

Sources of error were mostly identified in Chapter 3 and were discussed further in the previous points. Nonetheless, it is here reiterated that the predominant sources of errors may arise from:

- Navigational and ground truth samples position inaccuracies.
- Inherent noisiness of the backscatter data, including its inadequate reduction to estimates of seafloor backscatter only (i.e. filtering out unwanted variability i.e. cfr. Chapter 4).
- Challenges associated with the estimation of in-situ properties of sediment samples i.e. the validity of the ground truth data description approach.
- Representativeness of the samples i.e. challenges associated to understanding whether a single sample is representative of broader acoustic facies.

The survey azimuth dependence on backscatter is a potential candidate explaining sources of error on thematic accuracy, especially when dealing with multi-source multibeam datasets (often acquired in different orientations). While the angle dependence on the larger scale of a dunes flank is compensated in the backscatter imagery via the inclusion of the bathymetric reference model in the processing, ripples are not resolvable at the sonar beam footprint scale, and the various tilted facets influence the backscatter signal (Lurton et al., 2018). A possible example of this source of error is shown in Figure 6.6. However, the error observed could have originated from inadequately compensated bathymetry between disparate surveys since the effect of microtopography should have been counteracted by a backscatter compensation based on an angle of 45°. Generally, the backscatter dependency on microtopography, such as micro-oscillatory ripples, is manifested between 20° and 40° (Lurton et al., 2018, Montereale-Gavazzi et al. 2019), so that beyond this point the effect of sand ripples is erased by the mosaicking. From a gridded backscatter point of view, this is rather advantageous since the backscatter level will not depend on the survey azimuth (ideal in view of merging future datasets) and the same seafloor should provide the same average backscatter response, irrespective of the

ripples and heading orientation. On the contrary, it also represents a loss of information that could be of interest in other applications, beyond mapping sediment categories and interested with current/microtopography links. This compares to valuable remote sensing terrestrial applications in which satellite-borne radar monitoring of the sea-surface gives access to wave swell direction when measured from different heading angles (Al-Habasheneh et al., 2015). A similar approach is applicable to MBES backscatter where surveying at various heading angles could give evidence of the local presence of ripples, possibly giving access to their rough orientation.

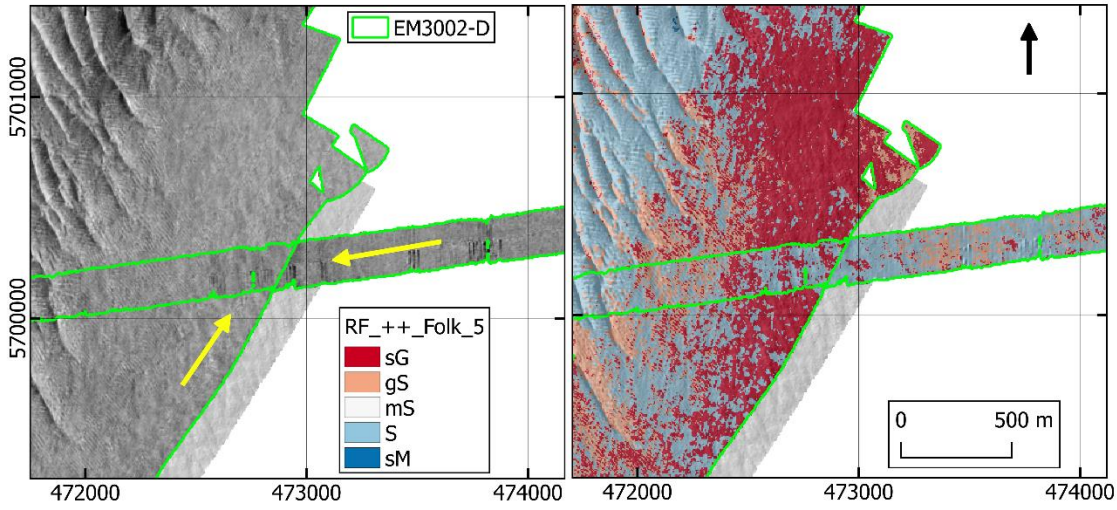


Figure 6.6 – Example of thematic error potentially due to; 1) survey azimuth dependence on backscatter acquired in different heading angles (yellow arrow) and at 300 kHz (for the same echosounder EM3002-D/ RV *Belgica*) and/or 2) artefacts due to relatively poor sea state conditions. The image is a detail of the Hinder Bank zone as classified in Chapter 3 using the $RF_{++FOLK(5)}$ model.

An important observation further enhancing the protocol of error estimation is the lack of spatio-temporal error propagation induced by non-complementary ground-truth (validating) samples as well their representativeness of the study areas assessed. The ground-truth data used in this thesis was always acquired in complement to the acoustic survey (i.e. within maximum 48 h of the acoustic data acquisition). By this approach, it was possible to generate error matrices that were representative of the entire classified area and as such, the accuracy metrics derived are highly trustworthy. This is a significant improvement compared to studies based on legacy or non-complementary validation data and which omit the notion of a dynamic seafloor, changing at various spatio-temporal scales (Stephens and Diesing, 2014; Mitchell et al., 2019). It must be noted that the sampling-delay tolerance will vary enormously depending on the study site (i.e. Chapter 4). Offshore and deep areas (unlikely affected by wave-induced sediment mobility) are rather stable compared to coastal and shallow areas which exhibit short-term variability at the scale of seconds and hours. For the latter scenario, a sample acquired at slack tide may not be representative of a backscatter dataset logged during flood tide.

The step further the matrix-based accuracy assessment, is a visual representation of model uncertainty. Besides the metrics, one has to consider the spatial component of the classification. In other words, how is the uncertainty of the model spatially distributed (on a per-pixel basis)? Based on the Random Forest classifier, instead of visualising only the predicted map (i.e. the aggregated majority votes of all trees in the forest with a given accuracy and associated metrics), one can map a measure of variability of the distribution of votes of single trees and visualise how its variability varies for each class and at different locations in the map. A composite image with a band per class is produced, each representing the fraction of trees in the forest that voted for a particular class. Using the Shannon formula of Entropy (for which Wegmann et al. (2016) is referred to) one can map by pixel the randomness in a sample of predictions (the tree votes) and appraise how consistent the classifier was at deciding upon class allocation. Figure 6.7 – exemplifies this on the $RF_{++FOLK\ (5)}$ model derived in Chapter 3.

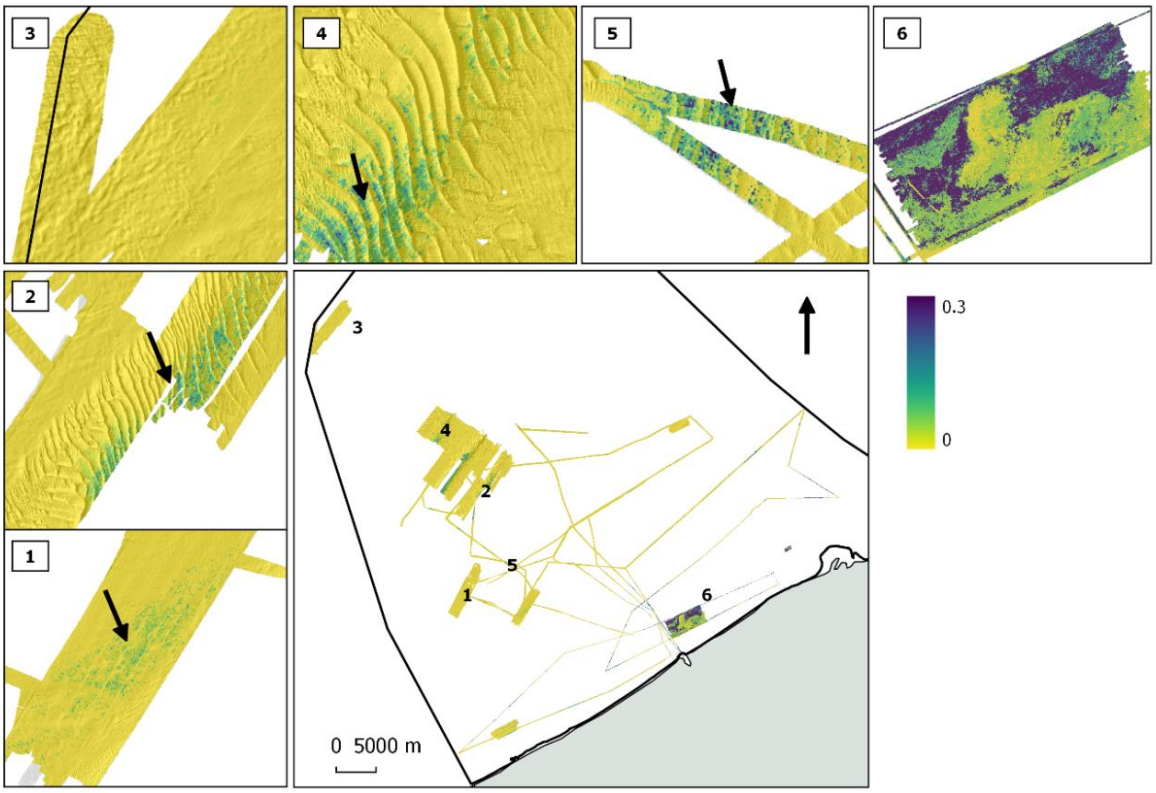


Figure 6.7 – Entropy of the classification results. The map analyses how consistent the classifier was at predicting classes at each pixel/location. It is important not to confuse this model-uncertainty output with thematic accuracy. Entropy ranges 0-1 (max 0.3 in this map). The map displays a generally strong certainty (low entropy; yellow-light-green). The highest uncertainty is generally distributed in the most sediment dynamic areas (high entropy; dark-green-blue) and results in transition zones between acoustic classes (1), at the top of the west and east Hinder Bank (2,4) and around the Ostend disposal site for the sM and mS classes (6).

6.4 Seafloor monitoring using MBES: variability and change detection

As defined by Goldsmith (2012) monitoring relates to the *"intermittent surveillance carried out in order to ascertain the extent of compliance with a predetermined standard or the degree of deviation from an expected norm"*. It is thus understood that a monitoring programme targets the detection of trends and/or changes from a "normal" or initial condition/standard over time. As discussed in the previous section, the "normal status" or standard condition, relates to knowledge of the system and an initial mapping effort, often comprising the strategic prioritisation of target study areas in respect to the near and far field of salient anthropogenic activities (Table 6.2, Fig. 3.1b, Chapter 3).

To be ecologically meaningful, a monitoring program must target the detection of physical conditions and processes liable to produce adverse effects on the benthic organisms (undoubtably a research objective paralleling the remote sensing change detection), thus altering ecosystem structure and functioning. Adverse effects can result from natural variability and anthropogenic activities and from a combination of both, occurring over short- (seconds to diel cycles) to long-term (seasonal to decadal) spatio-temporal scales (Halpern et al., 2015). Besides extreme geologic and/or atmospheric events such as storms, the scale and magnitude of modifications of the benthic substrate resulting from man's engineering and commercial activities in nearshore and offshore marine areas, can often and drastically exceed that of natural processes (Miller et al., 2002). Taking the example of dredging and disposal activities (Du Four and Van Lancker, 2008 - Figure 6.6A), beach nourishment (Hanson et al., 2002), marine aggregate extraction (Bellec et al., 2010), offshore aquaculture (Sutherland et al., 2007) or bottom trawling (Jones, 1992; Thrush and Dayton, 2002 - Figure 6.9 A-B) practices: significantly larger amounts of sediment are relocated in shorter time than would naturally occur. This reality is dictated by the design of these operations and is exacerbated by their logistics and contractual deadlines.

Considering that the benthic substrate type is a key driver of benthic biodiversity, dictating by large habitat type and suitability (Diaz et al., 2004; McArthur et al., 2010), its spatiotemporal assessment is considered as a primary requirement towards the implementation of sound marine management applications (Greene et al., 2008) and the links with multibeam technology and the applications presented in this thesis are obvious: information describing changes in the water-sediment interface takes us a step further by shedding light over the temporal dimension to our otherwise static perception of the seafloor. Developing the ability to spatially (continuous coverage of broad scales) and explicitly (at fine resolutions < 10 m) detect and understand seafloor changes is also highly pertinent in view of the global and sobering predictions of a rapidly increasing Blue Economy (i.e. the "urbanisation" of the marine environment) and of Climate Change (Halpern et al., 2015; Stock et al., 2018). As an example of a burgeoning Blue Economy, the

exploitation of marine aggregates in the BPNS started in 1976 with yearly harvests of approximately 30,000 m^3. Today, extracted volumes increased to the extent that overall (i.e. considering all extraction activities), over 3 million m^3 of aggregates have been harvested in a year time (Van Lancker et al., 2016), relocating it, in part, from the far-offshore Hinder Banks to the coastline. The strength and frequency of these operations can lead to drastic modifications of the seafloor geological (Virtasalo et al., 2018; Chapter 3) and biological integrity (De Backer et al., 2014; Rice et al., 2010), having the effect of altering the distribution and structure of benthic habitats, and potentially disrupting ecological functions which can have cascading repercussions on nature, economy and society (Gowdy and Mesner, 1998; Barbier et al., 2011). With respect to the globally changing climate, a key topic of research is, for example, the identification of spatial shifts in seascape structures such as habitat fragmentation and loss (Pittman et al., 2011). The global increase in coverage of environmentally warded areas has been achieved over the past century and continues to progress in several marine regions (Watson et al., 2014). However, despite the increasing pressures, and besides the Belgian case, wherein legally mandated long-term monitoring of the sand extraction is implemented by the Continental Shelf Service of Belgium (see Roche et al., 2017 for an overview), the application of management monitoring tools based on multibeam and backscatter data, has been generally limited.

This is inevitably associated with the novelty of the technology, survey-time needed for full-coverage mapping and the technicalities associated with controlling and automatically integrating the hydroacoustic and ground-truth measurements. Nonetheless, despite their paucity, change detection studies are increasing (e.g. Urgeles et al., 2002, van Rein et al., 2011; Rattray et al., 2013; Montereale-Gavazzi et al., 2017) as well as studies addressing variability and uncertainty of the backscatter measurements (e.g.; Madricardo et al., 2017; Gorska et al., 2018; Lurton et al., 2018; Malik et al., 2018; 2019). Cost-wise, the breakthrough these approaches can bring into assessing environmental status of the seafloor justifies the considerable costs associated with repeating surveys.

6.4.1 Environmental variability observation: the intrinsic and the unwanted.

The study of the short-term tidal environmental variability in Chapter 4 provided insightful empirical observations regarding the sensitivity of the backscatter measurements in the operational context (i.e. *in situ* where the mapping and monitoring are ultimately conducted), fulfilling the objectives set out under step 2. Firstly, it identified both unwanted and intrinsic types of variability. Then, it identified the sources, quantified the envelopes of variability over short-term spatio-temporal scales and investigated possible bypassing factors for the unwanted part of the variability (TL). Therefore, attention has to be placed on defining variability that is *unwanted* and variability that is *intrinsic* to a given habitat and hence cannot be bypassed. As we build knowledge towards the application of seafloor monitoring

using backscatter, or a combination of multibeam-derived spatial variables, this type of information is important to start constructing a baseline for both kinds of variability, to understand the measurements' sensitivities (both to intrinsic and unwanted) and to understand the short-term environmental cyclicity in the context of monitoring longer-term environmental and anthropogenic variability. For instance, in the context of monitoring before and after impacts of a given anthropogenic activity, Underwood (1994) put forward the strategy of monitoring a number of sites (small compared to the potential impact area) in order to estimate the magnitude of spatio-temporal variability for a given seafloor area and improve on the subsequent understanding of impacts. This type of information is also important to advance our ability to cope with the harsh environmental constraints. It needs highlighting that the maritime environment poses a considerable larger number of challenges compared to terrestrial remote sensing that is far less constrained by aspects of logistics.

Part of the research objectives set out under step 2 were addressed by investigating the effect of unwanted sources of environmental variability and how these potentially influenced the acoustic measurements. Attention was placed on quantifying the hydrological status, i.e. looking at mechanisms of acoustic energy dissipation in the propagation medium, possibly hindering the retrieval of the correct seafloor backscatter strength and having implications on the comparison of backscatter values of a same area during different dates with different hydrological conditions. This is homologous to the terrestrial remote sensing, where the atmospheric effects that modify the signal's amplitude and spectral characteristics have to be removed to retrieve correct signatures of Earth's surface reflectance (Gonima, 1993). Molecular and aerosol scattering, and water vapour absorption effects are directly comparable to the molecular relaxation processes which modulate sound absorption in seawater and to the presence of suspended sediments.

The two environmental sources of transmission loss looked at were the contributions of absorption due to seawater and that due to water-column and near-bed sediment load (Chapter 2 and 4). As expected, the contribution of seawater was negligible given both the shallow water and the year-round well-mixed water mass of the BPNS (van Leeuwen et al., 2015). In this scenario, the seasonal control of this environmental dependency can rely on near-surface values to satisfactorily correct the backscatter signal. On the contrary, preliminary estimation of the loss of acoustic energy due to suspended sediment presented in Chapter 4 (Exp. II and III, Table 5.8) showed that this effect can become important at high frequencies (300 kHz) in very shallow water (~15 m) and at a slant range of 70° (reaching 1 dB at 70° and ~ 0.5 dB at 45°). Previous studies concerned with sonar performance reported on dependencies of up to 3 dB for concentrations of the order of 0.2 kg m^{-3} in 100 m depth and at 100 kHz, pointing at the importance of also considering these properties of the propagation medium.

The biggest challenge here, relates to sample representativeness of the SPM concentration in the water body. While it is reasonable to posit that the seawater is a homogeneous fluid in the BPNS, and that the parameters modulating the seawater absorption are well-mixed in the medium, or at least differently yet homogenously distributed at various layers of the water column, suspended sediment exhibits rather complex patterns, especially in terms of the vertical distribution: reaching concentration peaks close to the seafloor, where the sediment transport is most prominent (Fettweis et al., 2006, 2009; Fettweis and Lee, 2017). This raises the following interesting observation: Assuming we are able to representatively sample the necessary parameters to compensate this effect (i.e. SPM concentration and grainsize), up until which interval (bin) of the sound propagation path is the signal-loss correction opportune? Consequently, this raises the conceptual observation pertaining to the definition of the water-sediment interface itself. The water-sediment interface targeted by acoustic classification is itself a "fuzzy boundary layer", referable to as a layer that includes the meter below and the meter above the interface (Anderson et al., 2008). Considering that high sediment transport and near-bed sediment concentrations are intrinsic to dynamic and nearshore seafloor environments, it is thus questionable to whether the highly concentrated part of the propagation path requires a filtering from the acoustic signal (in any case too short a path length to cause significant losses). Besides this, suspended sediments are known to exhibit complex vertical and horizontal patterns, orchestrated by currents, winds and wave action (Chen et al., 2010). This makes the sampling of this parameter at the scales and resolutions that would be required to correct MBES backscatter an overwhelming task, perhaps unrealistic by relying on current technologies (cf. Chapter 4). The latest technological advances in MBES water-column-imaging themselves would allow obtaining a spatially-explicit, continuous and three-dimensional understanding of the complexity of suspended sediment transport, though it remains highly challenging to quantify in terms of sediment concentrations (e.g. Best et al., 2010; Colbo et al., 2014; Kruss et al., 2015; Simmons et al., 2010, 2017).

As the seafloor mapping community moves increasingly towards data absolute calibration (Lurton and Lamarche, 2015), the need to control the environmental dependencies given by the operational environment will grow, requiring precise and explicit methods for proper compensation (i.e. de Campos Carvalho et al., 2013; Mayer, 2006). The case of suspended sediment however, at least from a monitoring viewpoint, may well remain as the case of cloudy and/or cloud-free satellite data in the terrestrial remote sensing realm (Figure 6.8).

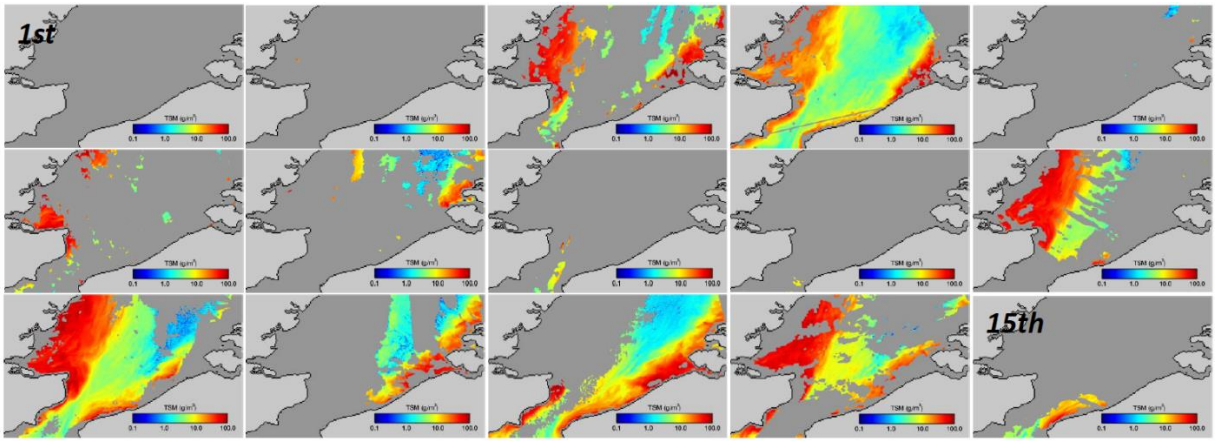


Figure 6.8 - Summary of the Total Suspended Mass (TSM concentration), during and after the experiment at the Westdiep site (Exp. II, Chapter 4). This was an attempt at retrieving finer scale sediment concentration values than extant models for this part of the North Sea. AQUA satellite images ranging from 1st to 15th of March 2016 (day of experiment was on 8th of March) and classified into Total Suspended Matter (TSM) expressed in g/l. Noticeably, due to cloud cover, no data were available for the day of the experiment. Satellite images are from the NASA AQUA in-orbit satellite and the inversion to TSM (in g/m^3, ranging 0.1 to 100 in the images above) following Nechad et al. (2010).

The SPM uncertainty estimation approach proposed in this thesis is by no means exhaustive, however our future ability to detect subtle changes may benefit from this kind of insights and empirical observations, possibly stimulating new research lines, such as the design of further experiments. This is especially pertinent in view of certain environmental monitoring applications, where some changes are likely going to be very subtle, though significant, hence requiring precise calibration and control of the measurements.

This viewpoint, wherein the expectation is that of comparing successive, calibrated, well-controlled and stable (on the short-term, at the scale of the measurement operation) measurements to infer changes in seafloor type, is a highly valuable ambition and investment, especially considering the advantages such a remote sensing approach would pose in the context of Rapid Environmental Assessment (Urgeles et al., 2002; Boyd et al., 2006; Sutherland et al., 2007; Siemes et al., 2008). This is a fully justifiable point of view: (1) reaching absolute calibration promotes global data comparability targeting harmonisation and the build-up of acoustic inventories (libraries) of backscatter signatures comparable in space and time (e.g. Fezzani and Berger, 2018, Chapter 3). (2) Controlling the measurements allows retrieving only the part of the echo that relates to the seafloor, omitting the external variance potentially embedded in it, and (3), under these premises, backscatter stability allows the repeatability of measurements: directly exploiting the "trustworthy" measurements and possibly implying a lesser sampling effort (as this ideal scenario is built towards in the long-term).

The conditional rationale behind this approach follows:

If multibeam backscatter signatures (grid ranges, average values or angular responses) can be confidently related to ***n*** seafloor type ***and*** repeated measurements are stable and accurate, ***then*** one can directly infer changes in seafloor composition.

However, this line of thought seems lawful in a “mapping for discovery” scenario, for example between neighbouring countries having to merge datasets, whereas in a monitoring context, it becomes a vicious circle since assuming that we are able to build a library of “acoustic diversity” once and use this information to compare it against future surveys, precludes the appearance of new features in the acoustic library (i.e. sediment types) of the seafloor area at stake. This lack of *a priori* knowledge is perhaps the greatest challenge faced by seafloor mappers, operating in an environment that is scarcely known. In fact, the point is often that of exploring and detecting novelties, especially in a monitoring scenario.

Indeed, while the previous conditional statement hides the body of knowledge that we need to build to disentangle the complexity of environmental variability, and possibly reach standards that will allow a more informative exploitation of the backscatter measurements, we also need to recognise the bypassing solutions formulated from decades of experience by the terrestrial remote sensing community, experiencing similar issues inherent to calibration, registration, rectification, geometric and radiometric reduction of the data (Singh, 1989; Coppin et al., 2004). Such issues have often been counteracted by post-classification change detection: a process by which classified seafloor cover distribution models are compared by matrix-based approaches (Chapter 5, and later reiterated).

Regarding the sensitivity of the backscatter measurements to intrinsic patterns of variability in soft sediment and highly dynamic areas (predominantly muddy - Exp. III), key findings of Chapter 4 identified the environmental source which shared the closest association with the backscatter measurements (and the bathymetric bottom detection) over a tidal cycle: a cyclical (slack-tide/semidiurnal) formation of ephemeral depositions of up to 30 cm of dense *“fluffy”* cohesive particulate matter, referred to as High Concentration Mud Suspension dynamics (Fettweis and Baeye 2015). On one side, this type of intrinsic variability has implications for the interpretation of high-resolution snapshot in time maps in a highly sediment dynamic environment such as the Belgian nearshore turbidity maximum area, posing the question how long a map is a valid representation for an area. On the other side, it also points at the sensitivity of the measurement to relatively subtle seafloor sediment changes which may be of great interest in other monitoring applications elsewhere. For example, a study by Sutherland et al. (2007), based on 300 kHz backscatter mapping, identified the relationship between waste material from a ceased fish farm activity and the backscatter data, capturing the extent of the impact and relating the backscatter intensity to a soft sediment patina of gel-mud, similar to the cohesive material sampled in the third experiment of Chapter 4 and frequently described by Fettweis and Baeye (2015 and previous works therein).

Considering that offshore fish-farming projects are planned in Belgium (Douvere et al., 2007) and the significant seafloor impacts of such activities on the seafloor are well-documented globally (Cook et al., 2006; Telfer et al., 2009), this suggests that a great deal of insights regarding mechanisms of waste material dispersal on the seafloor can be achieved in the framework of a long-term aquaculture-impact monitoring study based on these technology, and generally to the monitoring of anthropogenic deposits in the near field (i.e. where do the organic-rich sediment plumes generated by the windfarm piles (i.e. Vanhellemont and Ruddick, 2014; Baeye and Fettweis, 2015) end up and how do they affect the modelled 0.066 km^2 impacted area around each monopile?). Regarding the monitoring of far-field effects, it is critical that support systems and ecosystem models be developed, enabling to understand potential sink zones and gauge ACD applications.

6.4.2 Seafloor acoustic change detection

Change detection can be defined as: *"the process of identifying differences in the state of an object or phenomenon by observing it at different times"* (Singh, 1989). From the set of methodologies tested in this work there are a variety of possible approaches to seafloor change detection, each of which has advantages and disadvantages or is better suited to a given application: a method does not exclude the other and may even complement each other (take the example of detecting micro oscillatory ripples with AR and sediment type, irrespective of small-scale bathymetric relief, with CBI data-types respectively). Change detection based on multibeam backscatter, and associated data products, has been alluded to by numerous investigators (e.g. Zajac, 1999; van rein et al., 2011; Culloch et al., 2015; Snellen et al., 2018), though it remains a generally scarce application (Boyd et al., 2006; Du Four and Van Lancker, 2008; Roche et al., 2015), mostly due to the set of challenges previously mentioned, and particularly due to the relatively short time (approximately two decades) the technology has been put at the service of the scientific and maritime management applications concerned by the present study.

Research objectives set out under step 3 were addressed by testing the comparison of a suite of change detection methodologies; namely pre- and post-classification and an ensemble approach.

In Chapter 5 an analysis was tested, presenting a hybrid situation regarding backscatter data type used and change analysis applied. For example, Rattray et al. (2013) carried out a bi-temporal change detection analysis using the same approach as in Chapter 5 (*Post-classification*) though based on two independently derived models (using a Decision Tree classifier), each with complementary ground-truth data and accuracy metrics. Their study reported on considerable differences in backscatter data values between serial surveys due to a lack of system calibration and/or control on the measurements repeatability and stressed the severe implications this has on *pre-classification* change detection, where the data are

directly comparable (for a review see Singh, 1989). The dataset exploited in Rattray et al. (2013) could thus rely on post-classification change detection since data from two dates and sensors could be individually (independently) classified, minimising issues inherent to correction for hydrologic conditions and of acquisition and processing settings. For the backscatter time series dataset used in Chapter 4 ($n = 7$), only one complementary ground-truth survey was available. It is here that the careful control on the backscatter data repeatability allowed fully exploiting the entire dataset: a considerable advantage given costs associated with ground-truthing. In practice, it has been possible to produce an accurate model using supervised image classification, identify the dB ranges that defined the sediment type in the classification scene, and propagate this spatial and comparable information to the rest of the dataset for which there was a paucity of data for robust model training. This methodological framework showed to be highly informative, quantifying fine-scale seascape changes and identifying the dominant from-to transitions and signals of change at the level of the entire study area.

The promising application of such an approach can be exemplified by, for example, the need to monitor kelp macroalgae in temperate waters where rising ocean temperatures pose a severe threat to marine life. Kelp species provide the structural complexity needed by several marine species to thrive and control ecological processes of major importance (e.g. Wernberg et al., 2010). Several studies have demonstrated the potential given by MBES technology to map foundation species such as SAV (e.g. Kruss et al., 2008, 2011, 2015, 2017; McGonigle et al., 2011; Rattray et al. 2013). A further useful application of this approach would be in the Venice Lagoon, where considerable engineering modifications of the tidal inlets have been put in place and recently, Madricardo et al. (2017; 2019), conducted extensive multibeam mapping and monitoring. These approaches are thus transferrable to an array of marine environments and applications and can provide the means to quantify seafloor integrity in space and time.

What remains remarkably challenging from a change detection viewpoint is to, except from detecting the obvious impacts (i.e. Fig. 6.10), decompose the changes into natural and anthropogenic ones. It is clear from the application presented in Chapter 5 that, for this particular BPNS seafloor area, the most prominent signals of change were due to the morphological evolution such as dune migration: readily identifiable in the from-to transitions between substrate classes (i.e. sand/gravel) and in the persistence/gain/loss maps therein presented. Their interpretation provides useful insights; the “fine-to-coarse” gain signal within the gully part of the study area, was predominantly characterised by strong lineation indicative of the strong bottom currents and sediment transport, but possibly also due to frequent bottom trawling. Understanding whether this change was accelerated by sand suspended from extraction activities entering the system or from other sources and causes, remains hampered by a lack of knowledge of the natural variability. Indeed, the occurrence of

mobile sand (such as sand patches, dunes and/or ribbons) at the seafloor leads to an intrinsic natural variability that has implications for the design and interpretation of changes in repeated surveys. At the same time, for this particular case, knowledge of the epibenthic biota resilience in respect to natural sediment transport patterns, remains unexplored and requires a highly-multidisciplinary approach (i.e. laboratory measurements targeting species and macrobenthic communities - Zajac, 1999; Miller et al., 2002) to disentangle these intricate patterns (Figure. 6.9).

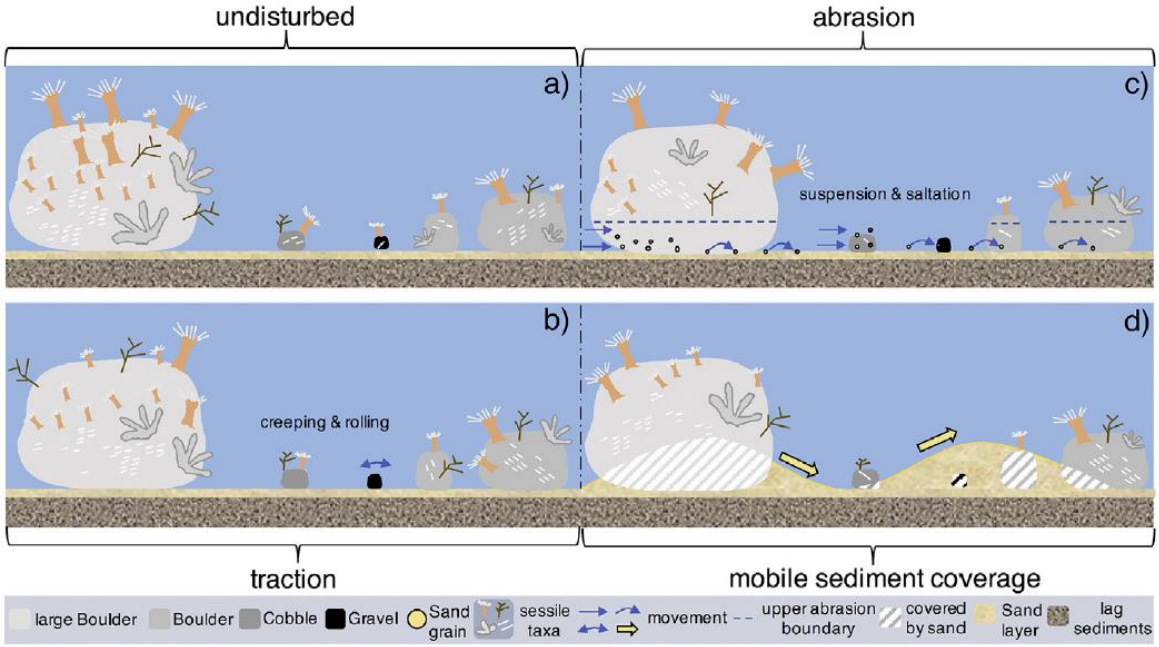


Figure 6.9 – A schematic model of the expected colonization of hard substrate features by sessile benthic organisms in respect to processes of abiotic disturbance. (a) Undisturbed conditions/high colonisation; (b) Traction of small substrate features (compared to boulders) remaining uncolonized; (c) Abrasion induced by sediment transport leads to a vertical colonization boundary and (d) sediment dislocation/migration buries and exposes boulders and stones. Here, colonisation can occur periodically (image from Michaelis et al., 2019)

The monitoring objectives, frequency, extent and time-budget availability together dictate the type of sensors, surveying times and analytical approaches employed (e.g. van Rein et al. 2011). Post-classification can be highly informative where large scale full-coverage is achievable (Strong, 2015). On the other hand, this may be hindered by cost and time-budgets, making other options more viable such as interpolating between spatially separated and classified survey tracks (Anderson et al. 2007). However, this would come at the cost of introducing ambiguity due to interpolation between adjacent surveys (Gaida et al. 2018) and losing the valuable patchy/heterogenous spatial information sought after. Indeed, the transect/trajectory surveying strategy devised in this thesis (Chapter 3) poses several advantages in view of future evaluation of from-to transitions between the main substrate classes (sensu Folk and MSFD). Repeating the suite of advanced acoustic seafloor classification and change detection applications tested throughout these investigations will enable, for example, the detection of a hypothesised westward expansion of the Mud class along the coastline in respect to dredging and disposal

activities. To provide broader spatial context to the reallocation (faith) of the sediment disposed of at the Ostend disposal site presented in Chapter 3, interpolation between MBES track-lines may better capture large scale, regional changes envisaged by the Belgian MSFD monitoring programme.

Both from a static and temporal acoustic classification point of view, the most prominent limitation is in the unambiguous allocation of coarse/heterogenous sediment classes to univocal acoustic classes. The masking/dispersion effect observed in Chapter 3 leading to class aggregation, and noted by several other investigators (Boyd et al., 2006; Buscombe et al., 2017; Gaida et al., 2018; Fogarin et al. 2019) including in similar (non-hydroacoustic) remote sensing applications (Peiman, 2011), is a serious matter and will inevitably hinder an effective change detection between heterogenous sediments. It is therefore crucial that post-classification change detection be interpreted by considering the class separation potential dictated by the "geoacoustic perception" achievable by the operating frequency in use and the consequent reduction of the information content to the attributable degree of generalisation (i.e. that of the classification scheme). This is where including morphometric derivatives and researching their influence on class prediction at various spatial scales, along with possible ways forward previously proposed (see *Ways forward*), may considerably enhance the detail achievable in terms of class separation based on backscatter (and a combination of acoustic spatial datasets) and hence, a more detailed and information-rich change detection.

It is also true that, as previously noted, categorical/thematic mapping may not be suitable especially for highly heterogeneous sediments. While constraining information into predefined boundaries is an appealing approach to the communication of information (i.e. in the science to managers and decision-makers interaction) from a scientific standpoint, this could be perceived as a loss of information and may hamper the gain of new knowledge. An example follows: while a 3 and 5-class model is highly informative now, it may not be so in future. There may occur changes that may pass unnoticed if the acoustic data were constrained (forced) into the same thematic classes to allow comparison of classified scenes in time. Expert interpretation remains of very high value. This is especially the case where the seafloor scene is of a highly heterogenous nature: i.e. gradational/ecotonal transitions that are of great interest in a monitoring (and ecological) context, may not be recognised by a hard/crisp classification. It is here that the advancing approaches in system's absolute calibration, paralleled by research of the environmental (and instrumental) variability, will together provide a more robust background to undertake pre-classification change detection. After all, mapping and monitoring without classification (see the unclassified data in Figure 6.10), could better reveal given patterns of environmental change when the scale of assessment requires it. While image-based classification is highly informative to provide area-size determination over large distances, and indeed to derive easy-to-communicate from-to transitions between main sediment categories as shown in

Chapter 5, pre-classification may show unexpected processes and trends that may otherwise be blurred by the classification (Hass et al., 2017). An interesting approach that deserves testing, would be to repeat the surveys presented in Chapter 3 and, assuming a good control of the repeatability of the measurements, make a direct comparison using simple image-differencing (or else). The change detection would then focus on analysing the persistence (around the 0 of difference image histogram) and analyse the bins in the tails of this distribution to observe fine-scale changes. Fuzzy classification could in turn be used to identify gradational patterns of change.

Change detection based on multibeam and ground-truth data is the inevitable evolution of static seafloor mapping and represents a scientific line of research that will undoubtably continue to be pursued by investigators and will thus continue to evolve (Anderson et al., 2008). The research herein presented provides sound and repeatable methodological frameworks for following investigations into mapping and monitoring seafloor substrate type (and beyond) in the framework of the Belgian Marine Strategy Framework Directive and in view of detecting changes in seascape structure and sediment type farther afield.

Undoubtably, to gain a synoptic, more holistic, understanding of the extent of the impacts caused by the spatial mosaic of anthropogenic disturbance on the seafloor, will require future research identifying the natural scale of variability, both within and between areas of different seafloor substrate types, where current and sediment transport are closely associated to the short-term (hours to seasons to years) fluctuation of the hydroacoustic measurements and to the intrinsic seafloor natural variability. There is no need to discuss how invaluable the improvement of extant current and sediment transport models would bring to an improved prediction of the spatial extent of anthropogenic impacts (e.g. van Lancker et al., 2016), selection of priority areas to study seafloor changes and generally, to the prediction of seafloor habitat *sensu stricto* (i.e. by including physically dynamic variables in the habitat classification process as for example in Rattray et al. (2015)). Increasing volumes of high-resolution data will progressively become available and enable the construction of models of high-resolution dynamic physical variables (current and sediment transport), whose predictions could be validated by remote sensing approaches; effectively cross-validating the two methods.

Furthermore, future in automation of remote sensing data acquisition through robotics, such as Autonomous Underwater Vehicles (AUV) will enable mapping of the entire seafloor at high resolutions (Mayer, 2006), while simultaneously gathering a range of oceanographic and videographic information. Evidence suggests that autonomous and unmanned technology is gaining momentum and is able to provide substantial improvements of cost-time budgets in large-scale monitoring efforts resulting in increasing temporal and spatial resolution of the acoustic, videographic and physico-chemical data required (Van Lancker and Baeye, 2015; Paull et al., 2018; Tillin et al., 2018; Jones et al., 2019; Zarayskaia et al., 2019). However, the operationalisation needs careful consideration with evaluation of the effort of

continuous and careful control of the vehicles, the required ground-truthing and performance in areas with higher water and sediment dynamics that largely impact on the survey and videographic quality.

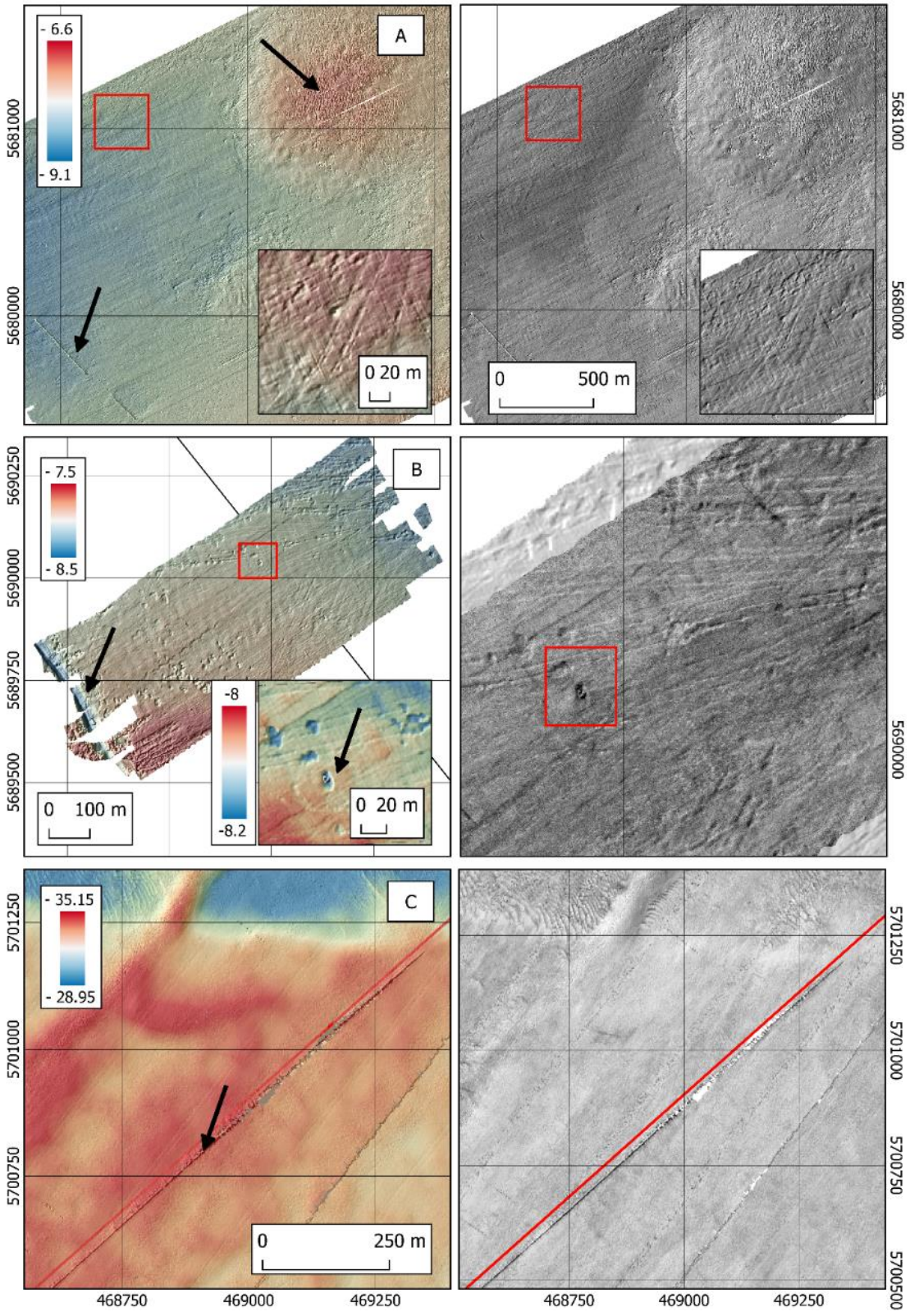


Figure – 6.10 - Detecting the "human footprint" on the seafloor. Caption continues on next page. Details of bathymetry and backscatter imagery acquired within three zones. Black arrows throughout

the images denote salient anthropogenic impacts. A; 0.5 m bathymetry and backscatter details of the Ostend disposal site study area (Ref. Figure 3.1b). Note the presence of a mound (a bathymetric difference of ~ 2.5 m compared to the surrounding) as a direct consequence of the disposal. Note the complex backscatter pattern showing high reflectivity, denoting sand/shell mixtures, over the mound and low reflectivity, denoting sludge/water-saturated mud, in the Eastern bathymetric low. It is hypothesised that the sludge, sand and shell material accumulating in the dredged channels of Blankenberge harbour, when deposited at the designated location, remain trapped on-site whereas the fine sludge accumulates in the bathymetric low (Van Lancker et al., 2007) and part of it reallocated in the adjacent areas depending on tide and current. A second arrow denotes a pipeline, likely reinforced by protective mats, resulting in higher reflectivity. Bottom trawl marks are widespread witnessing localised scouring and erosive processes. B; 0.25 m Bathymetric and backscatter details of the MOW 1 area (ref. Chapter 4). The arrow on the left-hand side denotes a dredged channel resulting in a ~ 1 m bathymetric difference. The second arrow in the subfigure denotes the position of the benthic lander that was deployed there, and that the campaign 17-322 was scheduled to recuperate following the end of the experiment (ref. Chapter 4 – Exp. III). Note the scouring around the tripod, as well as occurring in complement to the visible trawl marks and appearing as elliptical “pockets” ~20 cm deep. These may occur either due to trawling-current interactions or be due to previous deployments of the benthic lander. The backscatter image displays lower tones with them suggesting accumulation of finer sediment matter. Bottom trawls on the contrary result in stronger acoustic returns. C; 1 m Bathymetry and backscatter details of a part of the Hinder Bank gravel gully. The arrow denotes the passage of a pipeline.

6.4.3 Complexity of the problem, advancing technology and years to come

Over the past two decades, the underwater remote sensing technology has progressed at a fast pace, providing increasingly better ways of imaging and monitoring the hardly accessible seafloor (Mayer et al., 2018). Paralleling the technological advancement, standards in backscatter data acquisition, processing, calibration and interpretation (Lurton and Lamarche, 2015; Eleftherakis et al., 2018; Fezzani and Berger 2018; Malik et al., 2018b; Roche et al., 2018; Schimel et al., 2018; Weber et al., 2018), classification (Anderson et al., 2007, 2008), definition of uncertainty budgets (Fonseca et al., 2019; Malik et al., 2018, 2019) and ground-truthing the hydroacoustic data (e.g. Ferrini and Flood, 2006) have substantially matured, and considerable advances have been made since the first meeting of the BSWG-GEOHAB community. Particularly stimulating is the willingness of cooperation by sonar manufacturers, making the black-box character of the multibeam hard- and software increasingly more transparent and accessible (Schimel et al., 2018). These are clear indications that, similarly to our counterpart colleagues in the terrestrial remote sensing realm (e.g. Coppin et al., 2004) and despite the comparatively harsher operational environment, the applications of ASC and ACD will evolve towards well-established *modus operandis*. The research presented in this thesis provided ample evidence supporting the conclusion that acoustic seafloor classification and change detection can, under certain degrees of confidence, be employed to quantify spatial and temporal changes of interest for an array of applications. However, there are four main factors that remain very challenging in the field of hydroacoustic seafloor mapping and change detection and are briefly discussed hereafter. These refer to: 1) invest in methodologies that enhance the backscatter discrimination potential for heterogenous substrates; 2) increase survey coverage and frequency; 3) work towards standardisation and control of the measurements and 4) develop knowledge of the system, needed for tools supporting remote sensing such as ecosystem simulation models. The success of their evolution will depend on synergies and coordination between several interested parties and further technological/industrial and scientific developments.

In the previous section, "*How many classes? Which classification scheme to use? "What can my sonar see"?"*, ways forward and exciting new possibilities enhancing **classification performance** were put forward and discussed, collectively and partly addressing the fundamental issue of backscatter data **discrimination potential** at a given operating frequency and how the data be processed and analysed. When prescribing thematic classification schemes (such as those herein investigated), a degree of intraclass dispersion is expectable given certain grain size attributes (i.e. bio- and geo-clastic content) will dominate the backscattering response, resulting in a masking effect and leading to class aggregation.

Besides this, the complexity of the seafloor acoustic classification and change detection problem faces further challenges which are indeed the centre of attention of ongoing developments.

The need to increase **survey coverage** to achieve broad-scale datasets and effectively capture long-term modifications of a given study area (i.e. at the level of an EEZ) in respect to both naturally- and anthropogenically-induced variability, has equally been recognised for long by investigators on land and took several years to effectively harmonise and compile (i.e. Global Land Cover map - Gong et al., 2013; Hansen et al., 2013). In this regard, it is clear that the maritime remote sensing community faces logistical challenges that are not trivial in this respect and, as previously mentioned, will require a substantial improvement of automation of hydrographic operations; this must come from efforts in robotics and underwater communication and remote and repeated acquisition of the disparate datasets (semi)autonomously (e.g. gliders - Van Lancker and Baeye, 2015; mine countermeasure UAV fleets - Paull et al., 2018; Multiparametric autonomous vehicles - Jones et al., 2019; UAV fleets and remote data transmission - Zarayskaia et al., 2019). It must be stressed, that daily datasets are available on land (up to a spatial resolution of 61 cm at nadir for a Quickbird sensor, as an example amongst several, i.e. http://glcf.umd.edu/data/quickbird/) whereas due to the operational constraints of the oceanic environment, changes will have to rely on longer and less regular time-intervals between serial surveys (Chapter 5).

Backscatter **data comparability** is equally of paramount importance. Advancing research into backscatter measurements' stability control (e.g. Chapter 4 and 5, Roche et al., 2018), pragmatic and feasible absolute field-based calibration of sensors (e.g. Eleftherakis et al., 2018) and identifying the sources and magnitudes of instrumental and environmental variability (e.g. Lurton et al., 2018; Malik et al., 2018; Montereale Gavazzi et al., 2019), together promote standardisation. These are critical developments allowing compiling (merging) backscatter datasets into large-scale, transboundary, geographical coverage maps (e.g., as foreseen in the Seabed 2030 initiative; Mayer et al., 2018) and possibly a better utilisation of serial measurements.

Considering large-scale monitoring, for example at the level of the BPNS or at the scale of the Southern Bight of the North Sea, a better interpretation of the detected changes between available datasets would be achievable given a better **knowledge of the natural dynamics of the ecosystem**. From a terrestrial remote sensing perspective, knowledge of, for example canopy cover dynamics and forest systems, has been achieved by monitoring of the landscape attributes of variability over "*at least a few decades*" (Coppin et al., 2004; Ban and Yousif, 2016) and these patterns were studied in respect to the variability of the spectral measurements (similarly to the experiments presented in Chapter 4). This is where beginning the compilation of a baseline for the various kinds and scales of environmental variability (such as those observed in Chapters 4 and 5), quantifying the envelopes of variability, will provide invaluable input as we reach higher-level standards in harmonisation and operation of both measurements' acquisition, processing and correction, and of echosounder calibration.

Furthermore, at regional scales, effective monitoring is hindered by the paucity (if not complete lack) of **spatially-explicit models** (i.e. at the scales of interest; i.e. nested in the high-resolution MBES bathymetry) describing the natural fluctuation of dynamic variables. Besides being limited to a point in space, long-term timeseries of oceanographic parameters such as those collected by Fettweis and co-workers using the benthic lander presented in Chapter 4 (e.g. Fettweis et al., 2015) provide the means to appraise the magnitudes of variability in such an environment advancing significantly the system knowledge (here referring to the turbidity maximum area of the BPNS). Networks of marine observatories (i.e. benthic landers and semi-autonomous stations deployed on the medium to long-term) would inevitably allow a better appreciation of the scales and kind of environmental variability. These datasets contribute to our long-term capability of exploiting remote sensing for change detection, improving the integration of remote sensing with supporting methods such as ecosystem simulation models.

Detecting real changes against natural intrinsic variability, and against unwanted environmental variability influencing measurements, is a challenge the terrestrial scientific community already recognised for at least three decades (Singh 1989; Coppin et al., 2004). Chapter 4 demonstrated this interplay of variability and the difficulty this poses when interpreting natural form anthropic changes, especially in highly dynamic coastal environments. In most cases, except for areas with very-high sediment dynamics, detecting major and categorical changes; i.e. a conversion of seafloor cover resulting in the replacement of a given class by another one, will be readily apparent. On the contrary, more subtle changes (rather a *modification* than a complete *conversion* to another state), may pass unnoticed by a classification, unable to capture enough subtleties, especially a bi-temporal one, and will have to rely on a sufficiently long time-series dataset to study trends and oscillations of average backscatter strength (e.g. from ROIs) in respect to the surrounding ambient noise (as discussed in Chapter 4). This is because changes in seafloor composition can either be prominent (i.e. abrupt and drastic from-to changes) or can be rather subtle, slowly modifying the system and resulting in gradual modifications, such as accumulation of organic material (e.g. fish farm pen derived enrichment) or overtopping of gravel beds with fine sediments reaching far-field study areas (e.g. fine sediment outflow from dredging vessels). The latter case will require careful consideration of the surveying frequency in respect to the acoustical penetration expected into the sediment.

Finally, the emerging field of ACD needs some framing. This thesis has endeavoured in this challenging topic and has identified a series of important points which are hereafter reiterated. The following points (challenges and opportunities) are hoped to stimulate new research and experiments.

1) First of all, there is a great need of testing change detection methodological frameworks for which there is a paucity of application in the underwater remote sensing community. It is clear from the work herein presented that

both processing levels of backscatter data can be employed to gain an understanding of the seafloor in space and time at local scales (i.e. within circumscribed areas and mostly using AR), and at regional scales achieving the 'big picture" through CBI. Both pre- and post-classification change detection approaches have their merits, and it is advocated that comparative studies be commenced on available serial MBES datasets (an ongoing project at the level of the BSWG-VARIMONIT-GEOHAB). This is particularly important as the seafloor mapping discipline approaches the Big Data realm at a fast pace, with ever increasing volumes of data becoming available and being compiled by multiple initiatives. Testing change detection on regional datasets (i.e. at basin scale) remains highly challenging, and data compilation should consider seafloor dynamicity from the very start.

2) Besides the limitations of thematic classification identified in the previous section of the Discussion, it is worth arguing that especially regarding change detection, there exist the need to compare change detection based on both categorical and continuous classification of the remotely sensed data as well as on unclassified data. Early studies in the terrestrial remote sensing community too (Foody and Boyd, 1999) pointed at the fact that fuzzy classification (such as that tested and discussed by for example Lucieer and Lamarche, 2011) may provide richer information of the change: especially where seafloor changes occur at scales that are smaller than the spatial resolution of the echosounder in use; i.e. leading to "mixed pixels". This was especially evident in the previous section discussing the *Quantification of spatial uncertainty* where the RF categorical classification showed the highest uncertainty in the transitional areas between classes and over patchy/heterogenous areas within the overall predictive model. Consequently, it is critical to consider the sensitivity of habitat maps to spatial and thematic resolution. Highest resolution possible should always be favoured (since upscaling is feasible and downscaling is not). Continuous-type classification should be further explored, especially in view of detecting gradients and ecotonal patterns and how the expectable intra-class dispersion resulting from the contribution of e.g. sparse shell and gravel clastic material affects the prediction of continuous parameters (such as percent gravel).

3) Interpretation of change detection results needs consideration of sources of environmental variability. These can be either intrinsic (cannot be bypassed, hence characteristic of the seafloor area) and/or unwanted (should ideally be bypassed; they are equally characteristic of the hydrology of a given seafloor, but they can be referred to as "exogenous" (unwanted), thus radiometric discrepancies are best filtered/compensated between serial surveys). In this thesis, acoustical envelopes of variability across multiple insonification angles were quantified and successfully related to sources of variance that influence the measurements over tidal temporal scales. Ideally, the magnitude and

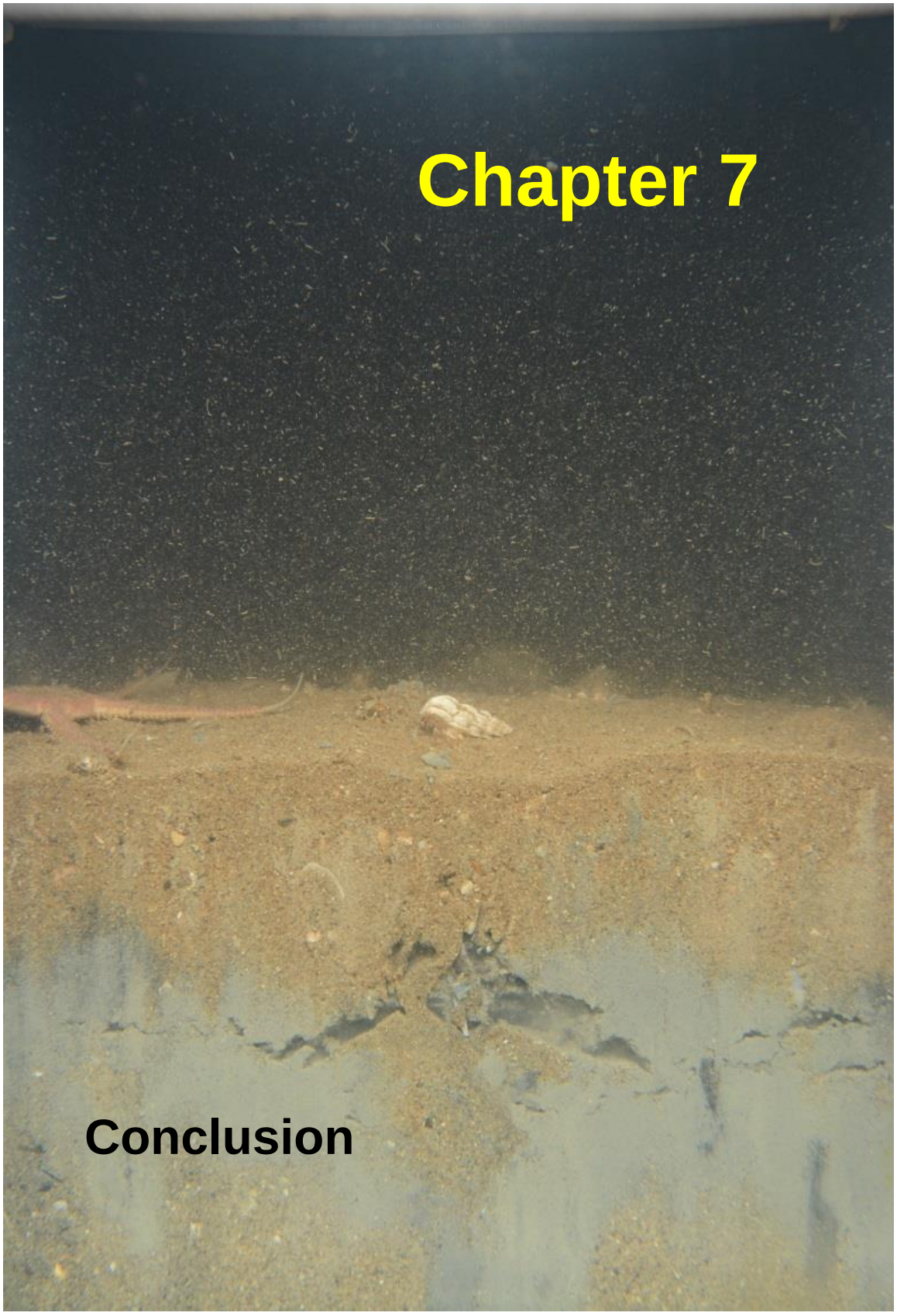
types of variability should be quantified and considered in medium- ad longer-term environmental monitoring. By way of short-term experiments, in different seafloor areas, such data can be obtained and will undoubtedly improve the understanding of the relative importance of different kinds of variability, and the effects this may have on the potential of discrimination between two categories. In this regard, it is key to understand both short-term cyclicity as well as the sensitivity of the seafloor backscatter measurements to a range of factors that could be of interest elsewhere.

4) Improving change detection by *a priori* filtering datasets is existing practice in terrestrial remote sensing studies, enabling direct comparison of the remotely sensed data. In the marine literature, a suite of radiometric corrections (e.g., in terms of suspended sediment concentration) can be applied, but further investigation is needed to incorporate this in backscatter acquisition, processing, interpretation and calibration standards. It is possible to estimate the error associated with hydrological conditions (seawater and suspended sediment absorption) and, assuming necessary parameters are sufficiently sampled, it is possible to compensate these effects in the seafloor echo of interest. However, representatively sampling the necessary parameters at the scales that influence the backscatter measurements remains very challenging. Similar issues are encountered in the terrestrial remote sensing community when comparing serial datasets acquired with different sun radiation angles (affecting the reflectance) and to the presence of clouds: simply making the data unfeasible for change detection. Here, understanding of the spatial and temporal scales over which the cyclicity of environmental phenomena (e.g., sediment transport) occurs is needed. This knowledge would inevitably allow a better planning of monitoring surveys for certain applications, for example prioritising surveying times and seasons in respect to given factors. This compares to, for example, selecting Earth surface images from the green season only to quantify changes in canopy cover (Coppin et al., 2004).

As a concluding remark, it is interesting to note that already in 1975, at the onset of terrestrial remote sensing using high-altitude spectral imagery, Aldrich (1975) proclaimed that the accuracy and stability of space-borne sensors could provide the means to monitor (automatically and accurately) the status of environmental disturbance (e.g., for forested landscapes). However, it took over three decades before this became a reality (Coppin et al., 2004). Regarding the maritime remote sensing, we have the same expectations that Aldrich (1975) envisaged on land, and it is promising to know that the underwater technology is increasingly more adapt to monitor and map ecosystem changes, building on several years of further exploring exciting methodologies. The research presented in this doctoral thesis is part of this progress.

6.4.4 Socio-political closure statement

Finally, while the MSFD provides a thorough legislative framework to manage and protect the European seafloor, strong institutional collaboration is required to set up and maintain a seafloor mapping strategy and monitoring program. Indeed, the project wherefrom this thesis took place, advocates to innovate in collaborative seafloor mapping by establishing a community of practice involving the main mapping institutions in Belgium (i.e. the Operational Directorate of Natural Environments of the Royal Belgian Institute of Natural Sciences, the Continental Shelf Service of the Federal Public Service Economy, Flanders Hydrography, Flanders Marine Institute and Belgian Navy). While this is ongoing, and active participation, exchange of ideas, methodologies and approaches occur, partly under a *Memorandum of Understanding* agreement and good will, a higher level of coordination between federal and non-federal departments, as well as research teams (i.e. from geophysics to geology and biology) would result in a more stringent monitoring of the marine environment, its drivers and pressures. Now that good practice guidelines and standards are in place, (e.g. IHO, 2008; Lurton and Lamarche, 2015), and ASC approaches have matured sufficiently, a new era is reached in which multibeam-based multi-parameter habitat mapping should be underlying the further exploitation of the marine environment. A nationally funded mapping programme could facilitate this, governed around issues and challenges that are nationally and internationally crosscutting. Last, but not least a good mapping strategy requires a truly multidisciplinary approach, with specialised research teams working together at sea and behind the desk, implying significant human resources, scientific and managerial personnel with state-of-the-art ecosystem and data-driven skills needed to implement the kind of current policies and make the most of the already available data. The approaches and tools developed and tested in this thesis demonstrate that there is great potential for a successful national seafloor mapping programme since the likelihood of success in achieving institutional collaborations for the production of high-quality benthic substrate and habitat maps, and their application to integrated marine management, is indeed high. This implies that the return on an eventual national (governmental) investment of resources would equally be high.



A Sediment Profile Imagery sample acquired in the Nearshore BPNS.

7. Conclusion

At the onset of sounding, mankind had only begun to scratch the surface of what expectedly turned out to be one of the most complex and enthralling (eco)systems, characterised by a plethora of interconnected properties and processes at multiple spatial scales, all having fundamental planetary atmospheric and life-controlling functions: The Ocean. As for man's awe and thrill for the Universe, the vastity and inaccessibility of the Ocean, and its mostly enigmatic and hidden floor, has long intrigued us, challenging our innate taste for exploration and discovery.

Over the past few years, there has been a remarkable breakthrough in our ability to map and visualise the seafloor, building an increasingly more detailed (patchy) picture. These innovations have come about through a variety of concurrent technologies, though the specific design of multibeam echosounder sonar systems contributed the most to our understanding of the seafloor and its composition beyond safe diving depth (≈30–40m), and considerably improved time and cost budgets of hydrographic operations. Embedded in the same seafloor-returning acoustical echo, over broad scales and at a very-high sampling density, bathymetry and backscatter together with ground truthing, and integrated via automated algorithms, begin to convey a degree of information that underpins the success of manifold applications and shows promising results in view of acoustic seafloor classification (ASC) entering the Big Data realm.

Thanks to these innovations, our knowledge of the ocean steadily grows, but so does the notion of its fragility and exposure to the turmoil of social and commercial activities which thus far, were recognised as mostly threatening our planet above sea level. From a bioeconomic and societal viewpoint (following Georgescu Roegen's line of thought - Gowdy and Mesner, 1998), the rate at which we adversely interfere (Callaway et al., 2007; Woodall et al., 2014; Halpern et al., 2015; Tekman et al., 2017) with an environment that we still poorly understand, shall be compared to the degradation rates as observed on land (e.g., the Amazonian rainforest, Scarrow, 2019). As with Earth observation, the enlarging and transdisciplinary field of ASC plays a central role towards our detailed and spatially continuous understanding of the ocean floor.

Observing Earth using in-orbit sensors is not trivial, though terrestrial remote sensing benefits from decades of engineering and academic experience, facilitated by a comparatively less constraining environment and by considerably greater fields of view and surveying frequency. Despite the tremendous improvements, underwater remote sensing faces considerably harsher challenges, starting from the complexity of investigating a target that is "listened to" rather than being "directly observed". As a fact, seafloor mappers deal with an environment that is mostly unmapped and unobserved, and generally have little access to *a priori* information supporting the investigations, and it is ordinary practice to acquire the various datasets *ad hoc*:

reason why a general trend towards automation, both at the acquisition (robotics/underwater communication) and processing (integration and classification) phases, is providing the necessary impetus for such a discipline to become an integrated part of our ongoing planetary exploration.

It is clear that exciting new approaches transforming hydroacoustic and ground-truth datasets into meaningful products that summarise reality (habitat, morphology and substrate) will continue to evolve and will be especially needed when programmes such as the completion of a geomorphologically-explicit global chart of the ocean floor, as foreseen by 2030 (Mayer et al., 2018), will be accomplished. Meanwhile, the ability to exploit the hydroacoustical seafloor backscattering phenomenon in terms of maximising the discrimination of benthic substrates and habitats is increasingly performant, and global standards for the acquisition, processing and interpretation are maturing (Lurton and Lamarche, 2015).

This doctoral research contributes to this stream of growing knowledge and specifically addressed and achieved the set-up of a baseline mapping effort towards the implementation of the European Marine Strategy Framework Directive (MSFD, 2008/56/EC) in Belgian waters, advancing the long-term, site-specific and regional monitoring of seafloor integrity (MSFD Descriptor 6). The state-of-the-art ASC and supervised and unsupervised data-integration routines allowed the production of accurate, repeatable, transferrable and spatially-explicit models of the seafloor nature, maximising the information content achievable from multibeam bathymetry, backscatter and their derivatives and identifying the salient limitations of prescribing seafloor classification schemes that were not designed with remote sensing in mind. The discrimination potential in respect to coarse and heterogenous benthic substrates was investigated in light of the available ground truth data and important implications on the assignment of thematic class to “acoustic diversity” (describing thematic resolution) were discussed. There exists a plethora of exciting and innovative classification methodologies, features, data-structures and technologies that are gaining momentum and are expected to considerably improve the current status of acoustic seafloor classification. However, while the remote sensing technology has drastically evolved over the past decade or so, ground truthing has remained relatively undeveloped and improvement of current (mostly mono-parametric) gears, gauging the characterisation of hydroacoustic field measurements, will be a critical research and technological investment. This investigation exemplified the utility of categorical mapping, and it identified and communicated important decisions on the spatial and thematic representation of seafloor habitat encountered in the map-building process, appraising their implications on ASC performance and applications. This is especially relevant as the research line of acoustic change detection is paved, denoting a new technological era in environmental monitoring.

Acoustic change detection research focused on developing knowledge of environmental variability. This is important to improve the understanding of environmental dynamics (over multiple spatiotemporal scales) and the interpretation

of static and serial MBES backscatter datasets. Dedicated field experiments were designed, beginning the construction of a baseline to quantify and discern between the intrinsic and unwanted types of variability that significantly influence the serial measurements and can have implications on data interpretation. These experiments, targeting the short-term (half-diel/tidal cycle) variability, recognise in full the dynamic character of the seafloor, often statically perceived, and open an innovative perspective on the sources and magnitude of the environmental variability to be expected in the operational environment over such short-term spatiotemporal scales. These sources of variability, and the kind of experiments, possibly improve the understanding of the measurement's sensitivities to an array of factors that are either scientifically observable quantities of interest, or unwanted and exogenous contributions that require filtering. It is clear from the experiments herein conducted that factors relating to the water-column (especially SPM concentration), to near-bed sediment transport and to seafloor-target geometry, can influence the backscatter measurements significantly in respect to the expected accuracy, and will thus have to be considered in any ACD application. Both kind of observations remain relatively scarce and require investigation considering the advantages of rectification with increasing potential for echosounder system calibration, the considerably maturing standards in acquisition and processing and the potential offered by direct comparison of backscatter measurements for change detection (but also in space, across systems, platforms and research teams).

Finally, acoustic change detection procedures were developed as a first and innovative critical step to assess and understand the evolution in environmental status of the seafloor, and methodologies were tested and evaluated that allowed quantifying interesting signals of seafloor change: the first key step enabling to relate patterns to causal factors. Deriving categorical patterns and trends of persistence and from-to transitions from multibeam acoustic imagery aids to decipher naturally- from anthropogenically-induced sediment dynamics and is pivotal in the design of monitoring surveys. This research aspect tested some principal change detection approaches developed in the terrestrial literature, proposing both use of unclassified and classified MBES datasets. This innovative line of research is important to appraise which methodologies can be employed on which kind of data-type (e.g. whether stable and relatively calibrated, fully calibrated or else), which patterns of change can be estimated, at which resolutions and over which spatial scales, and in a second, or synchronous phase, relate the seafloor change patterns to ecological trends, interpreted in the framework of natural and anthropogenic disturbance. Importantly as the way towards innovative monitoring of the marine environment is developed along with emerging and supporting technologies and tools, this research is required to develop the ability to match change detection phenomena with appropriate change detection (monitoring) applications.

7.1 Key research findings and challenges

Hereafter, the key research findings and identified challenges are concisely reiterated and grouped by research chapter. The reader is referred to Chapter 1b to consult the proposed objectives and research questions.

Chapter 3 - Brief

This research chapter endeavoured setting up a seafloor mapping strategy targeting the production of a seamless backscatter and bathymetry dataset allowing continuous prediction of the seafloor substrate distribution, by means of automated image-classification algorithms. In doing so, the chapter examined the validity of sediment-acoustic relationships in a field/operational setting and was therefore intended to provide researchers and end users a realistic point of view on what the MBES acoustic data can represent in terms of predicting material properties of the seafloor using conventional ground-truthing approaches. Following, the research endeavoured exploring the challenges and trade-offs associated with the pipeline of automatic seafloor classification and thematic mapping of seafloor substrate type. Considering the overall geophysical and ground-truth data acquired in the framework of this research chapter, the following key research findings were identified:

Key research findings of Chapter 3

- An important achievement of this research chapter is the demonstration of a pragmatic field-based-solution (stable and monitored at-sea reference area) to merge seamlessly disparate MBES backscatter datasets: a global challenge faced by the seafloor mapping scientific community.
- Regarding the previous point, an interesting research finding is the fact that where compensated backscatter imagery is corrected for the angular dependence using angles (or average values from a range of angles) beyond 40°, the effect of sub-beam-topographic-roughness polarization will be cancelled out, allowing seamless merging of sediment-type datasets acquired in disparate azimuthal orientations.
- Using the available ground-truth data, and based on exploratory data analysis, a number of insights were gained in sediment-acoustic relationships. At the level of the sample *loci*, moderate to strong univariate associations (*sensu* R^2) were found between backscatter intensity and the percent weight of individual grain-size fractions, within mostly heterogenous substrate types; and median grain-size diameter (D50), within relatively homogenous and unimodal siliciclastic substrate types). For the entire study area (i.e. the overall merged and seamless survey), moderate to strong associations were found by the multivariate statistical analysis, as well as when considering each study area in isolation. This suggests that different sediment parameters explain the backscatter collected at different locations.

- Importantly, it is observed that while Folk classes are a good global descriptor of backscatter variability, a strong degree of dispersion (in terms of backscatter values and basic statistics) exists for heterogenous sediment classes causing the reduction of the information content (by class amalgamation), and the subsequent generalisation of the depiction of the seafloor's spatial structure, achievable by thematic classification using geologically-conceived classification schemes (here referring to Folk, 1954 and from there originated, EUNIS classification).
- Here, a clear trade-off between backscatter discrimination potential (dictated by frequency) and sediment classification scheme, and thematic accuracy and resolution, was identified, shedding novel insights into the future research objectives and steps that have to be taken in order to improve this current limitation (see Chapter 6 for a detailed discussion).
- In the absence of a multi-parametric ground-truth sample description (first step in the classification process as described and visualised in the first box of Figure 2.14, Chapter 2), statistically-relevant geomorphometric variables were found to significantly improve the statistical and spatial accuracy of the modelled sediment classes.
- Comparing unsupervised (partitive clustering classification) and supervised (tree-based machine learning classification), it was found that the latter supersedes the former in all aspects when considering the "goodness of mapping"; i.e. thematic accuracy, spatial uncertainty, relevance of the contributing variables and validity of the geo-sedimentological patterns depicted in the final product.
- Lastly, clear trade-offs between number of sediment classes and scheme and thematic accuracy were detected, providing important considerations that can be of interest to seafloor mappers farther afield.

Main research challenges of Chapter 3

- Key challenges identified refer to those associated with the estimation of in-situ properties of sediment samples, including the precision of the positional accuracies. The representativeness of the ground-truth data description, being at the very base of the classification process, dictates the success of all subsequent operations (i.e. training the algorithms, computing statistical accuracies). It is therefore proposed that: 1) novel multi-parametric ground-truthing gears be designed and tested, and 2) extensive review studies be set-up, targeting the compilation of empirical data (regarding sediment type, variables and coincident backscatter intensities) available from the published literature. Such efforts would inevitably improve the way in which ground-truth samples are described, for which no consensus currently exists (e.g. from visual qualitative observations to geotechnical analyses of the sediment).

Chapter 4 - Brief

This highly-experimental research chapter endeavoured studying, observing and quantifying seafloor MBES backscatter variability for different seafloor areas that is due to short-term environmental cyclicity (i.e. tidal cycles). This research was intended to identify the sources and magnitudes of variability and therefore to provide surveyors and end-users with an improved understanding of how data, recorded *in situ*, be affected by such factors and subsequently, how to identify and deal with unwanted (external, to be filtered out) and/or intrinsic (characteristic of a given seafloor setting) sources of variability. Furthermore, the research provides important insights on how to set-up such experiments, highly-relevant to the utilisation of seafloor MBES backscatter in the operational environment, where environmental monitoring is ultimately targeted. Understanding how the environment influences the measurements against the resolution needed to detect true seafloor changes, is a critical first step towards the implementation of monitoring strategies that use such a technology.

Key research findings of Chapter 4

- The research detailed an experimental set-up needed to quantify sources of environmental variability, providing a solid basis to conduct future experiments within predominantly muddy, sandy and gravelly seafloors.
- Similarly to the previous chapter, this research demonstrated how standardising operational procedures, in terms of acquisition and processing, allows comparability, and therefore a better exploitation of repeated measurements, particularly in view of absolute system's calibration.
- The analyses concluded that different seafloor and hydrodynamic settings vary considerably differently, and the backscatter measurements therein logged accordingly.
- The sources of variability identified refer to: polarization of sub-beam topographic roughness, hydrological conditions of the water medium (i.e. presence of suspended particulate matter and of salinity and temperature gradients) and seafloor mobility (i.e. near-bed sediment transport, processes of cyclical erosion/deposition).
- With regard to bypassing and/or correcting for the identified variability, methodologies have been implemented that allow the quantification of Transmission Losses, necessary to reduce the backscatter values to estimates that reflect the seafloor as oppositely to other processes (i.e. processes that need to be excluded when applying Acoustic Seafloor Classification and/or Change Detection).
- This research chapter endeavoured identifying and discussing the implications of short-term variability on the use of MBES-measured BS for longer-term monitoring and whether such variability can hinder the detection of real seafloor changes by the backscatter measurement proxy-approach.
- The most prominent implications are: tidal periodicity and seasonality calling for careful consideration, especially in shallow areas with soft-material sediments

and high sedimentary dynamics. Indeed, successive surveys of a same area may provide different information at various time scales (from day to year). In this regard, it is important that the tidal dependence is analysed per MBES-BS time series.

- In a change detection framework using backscatter only and based on small ROIs, spotting outliers (i.e. abrupt changes in sediment response) will be relatively straightforward in the clear water and stationary areas since the magnitude of the short-term variance remains within the envelope of sensor sensitivity. On the contrary, the intrinsic "noisiness" (i.e. periodical variability) of the nearshore areas results in a potentially masking/blurring effect of changes in seabed type, introducing uncertainties due to the status of the water column (i.e. turbidity) or to the "mobility" of the water-sediment interface.
- Due to this, within such areas, the stability threshold must be defined contextually in accordance to the governing sedimentary environment, and a transition in seafloor status can only be detected from a trend analysis on a sufficient number of serial surveys. Direction and consistency of the trend, regardless of the noise envelope, can be a valuable proxy of change bypassing conflicting results from surveys acquired at different tidal and/or seasonal moments.
- The experiments demonstrated the sensitivity of seafloor backscatter to subtle seafloor changes that may be of interest in other applications, for example in monitoring sludge dispersal in respect to dredging and disposal sites, fish-farms and installation of anthropogenic infrastructures.

Main research challenges of Chapter 4

- An important challenge is the spatio-temporal resolution of the multi-parametric sampling efforts. In order for the disparate datasets to be closely comparable, samples should be acquired at the exact same moment and the exact same location. As this was not possible with the set-up detailed in Chapter 4, the "by-average" approach was pursued, homogenising the datasets to a uniform average time stamp (the mean time of acquisition within a pre-defined survey region). Future experiments should carefully plan the sampling frequency of the various instruments and work towards the improvement of the current set-up, for example by having benthic-lander-mounted transducers and diver-collected seafloor samples (improving their positioning in respect to the seafloor acoustic samples).
- Representativeness of the sampled parameters: this challenge relates to the difficulties associated with representatively sampling the parameters needed to study the acoustic variability. Suspended sediment in particular, shows high spatio-temporal variability, making its sampling highly challenging. It is therefore reiterated that coincident MBES water-column data be recorded and studied in conjunction to the seafloor serial backscatter dataset.

Chapter 5 - Brief

This research chapter endeavoured setting up acoustic change detection methodologies dedicated to improving the quantification of the seafloor's dynamic character, as well as to test methodologies applicable in the context of environmental monitoring in respect to anthropogenic activities. The research was intended to provide surveyors and users with a set of repeatable approaches to quantify seafloor changes.

Key research findings of Chapter 5

- Stable and repeatable backscatter serial datasets acquired within low-dynamic seafloor environments allow an effective change detection.
- Where a paucity of samples exists for the entire MBES time-series dataset, but sufficient data are available for one single dataset, and where rigorous data acquisition and processing standards have been employed, the supervised and accurate information identified in one survey, can be confidently extended to the remainder of the time-series, allowing its full exploitation.
- The change detection methods applied showed that different change patterns of interest can be observed and quantified. Pre-classification was used to study trends within well-defined portions of the seafloor (similarly to Chapter 4), whereas post-classification proved very useful to understand the broader picture: i.e. that of the entire study area.
- Post-classification is particularly recommended where issues of data rectification arise, by allowing the relative comparison of disparate datasets (e.g. the geographic delineations between 100 and 300 kHz datasets). Furthermore, this approach allows capturing important signals of change such as gross and net gains and losses, persistence and ratios to loss/gain of specific classes of interest.

Main research challenges of Chapter 5

- Natural from anthropogenic change remains highly challenging to quantify. To improve this understanding, better knowledge of the system is needed. Ideally, this is based on long-term time-series of measurements at key locations, as well as on validated high-resolution numerical sediment transport models (i.e. models that compare to the geophysical data resolution).

7.2 Contribution of the research to process knowledge

Finally, it is important to highlight that the methodologies developed and tested in this doctoral research contribute significantly to process-knowledge in the broader field of marine sciences, particularly w.r.t. the understanding of small and large-scale physical, ecological and anthropogenic processes driving marine ecosystems.

From a habitat characterisation perspective, it is clear that seafloor classifications, by integrating multibeam and ground-truth data, provide a synthesis of geomorphological and sedimentological attributes that are key drivers of marine life. Furthermore, spatially-explicit mapping allows the inference of physical processes that shape the attributes and habitat features. Particularly, this is the case for hard substrata which are often at the centre of attention of ongoing conservation and habitat restoration initiatives. Gravel beds provide small-scale structural complexity to fauna and are therefore pivotal for ecological processes of colonisation and succession (e.g. McArthur et al., 2010), or benthopelagic coupling. Geomorphology can also provide shelter to gravel bed epifaunal communities (i.e., the gravel *refugia* of Chapter 5) allowing these areas to evolve into hotspots of biodiversity within intensely fished areas.

Spatially-explicit seafloor mapping is also pivotal in sediment transport studies. As an example, the relative importance of hydrodynamics (tidal forcing, wave dynamics etc.) in an area becomes evident from detailed seafloor mapping, with derived sediment maps providing a proxy of the nature and availability of sediment, and bedform geometries providing insights into sediment transport pathways, all contributing to the validation of sediment transport models which still need tuning at the small-scale. Furthermore, the ephemeral patterns of erosion and deposition identified in Chapter 4, could feed into models that aim to predict the biological response to sediment budgets (as detailed in Miller et al., 2002).

Sound input of seafloor maps into numerical modelling tools at the small-scale are also critical for prediction of anthropogenic disturbance. As an example, the faith of sediment plumes caused by the marine aggregate extraction industry could be better predicted with small-scale data on sediments and geomorphology. Furthermore, subsequent sediment changes could be validated by well-controlled backscatter measurements (Chapter 4). The same applies to the validation of sedimentary change predictions as a result of the installation of offshore windfarms or other anthropogenic activities, such as disposal of dredged material (see Ostend study area, Chapter 3) that can modify current, sediment and ecological dynamics in the near- and far-field of human activities.

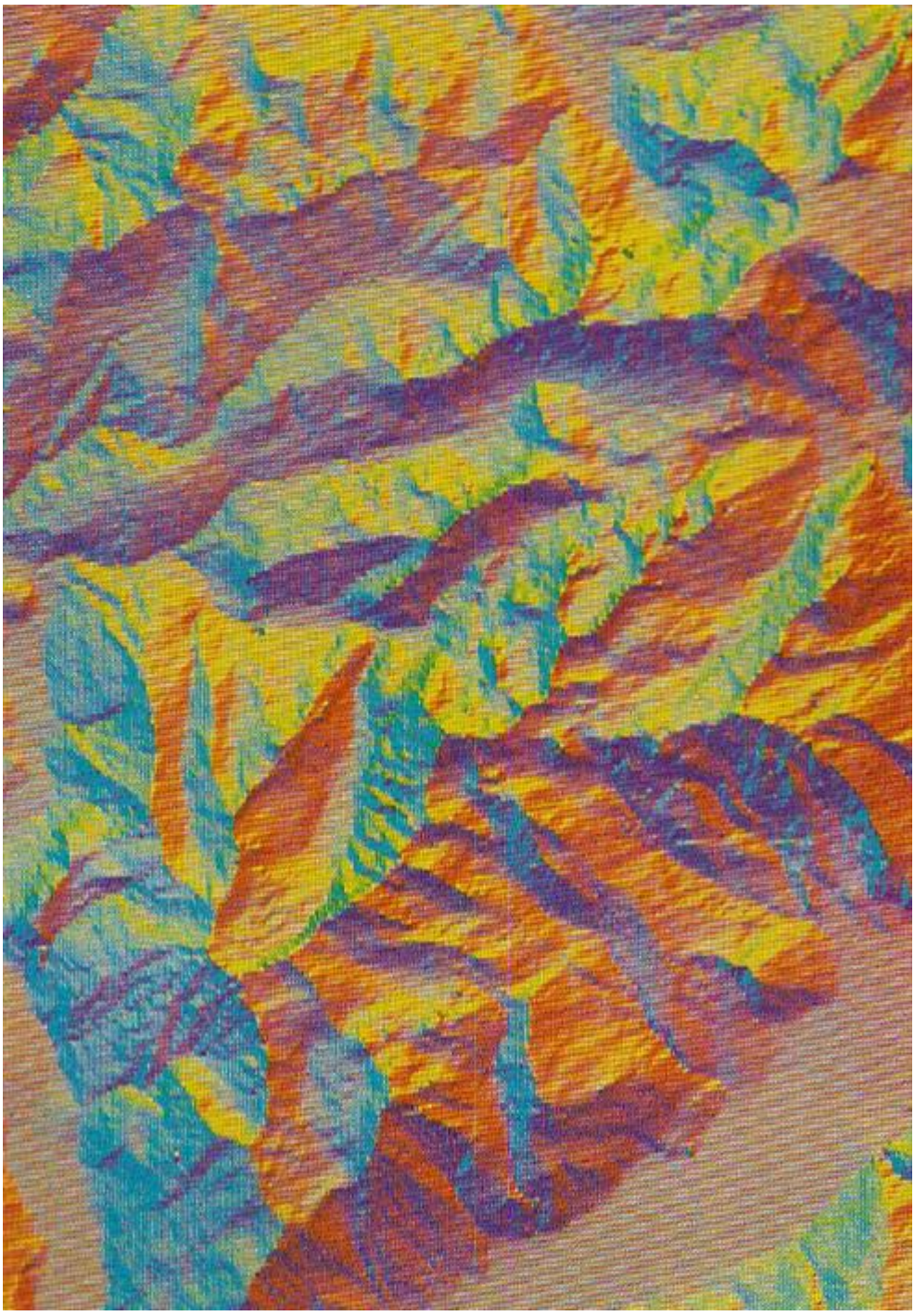
ACD methodologies (Chapter 5) are particularly powerful in providing direct evidence of seafloor change, for example the bi-temporal post-classification change detection clarifying the main sediment transport pathways of a sand bank/gravel gully system, and indirectly nutrient transport pathways too. This is highly relevant for the

identification of more or less suitable sites for epibenthic growth, hence critical to advice on the designation of MPAs and to support Environmental Impact Assessments. Furthermore, setting up monitoring protocols and quality control of repeated measurements should be seen as a fundamental aspect in developing the ability to distinguish between natural processes (e.g. short-term seafloor variations intrinsic to a given seafloor type) and “real” changes induced by anthropogenic interference.

Finally, it must be mentioned that this thesis contributes also to the study of the acoustic properties of seafloor sediments, supporting empirically the theoretical body of knowledge on geo-acoustic research (e.g. Hamilton, 1980). The work herein presented, in particular the experiments set out in Chapter 4, show how small-scale processes such as the composition, formation and short-term temporal behaviour of ripple morphologies influence the geo-acoustic response; processes whose understanding is generally based on laboratory studies.

Appendix E

Brief recapitulation of methods used in this thesis.



Coloured Hill-shading of a DTM. From Horn (1981).

A brief overview of the theoretical background to the techniques used throughout Chapters 2, 3, 4, 5 and 6 of this thesis follows.

Accuracy metrics from the confusion matrix

K statistic (Cohen, 1960). Formulae are from Banko (1998) given the exhaustive review presented therein.

The K coefficient measures the proportion of agreement after chance agreements have been removed from considerations; therefore, taking in considerations the off diagonal marginals as oppositely to only the diagonal entries denoting the correctly allocated pixels.

$$K = \frac{N \sum_{i=1}^{r} X_{ii} - \sum_{i=1}^{r} X_{i+} X_{+i}}{N^2 - \sum_{i=1}^{r} X_{i+} X_{+i}}$$

where

r = number of rows and columns in error matrix

N = total number of observations

X_{ii} = observation in row i and column i

X_{i+} = marginal total of row i, and

X_{+i} = marginal total of column i

The formula is better interpreted by the following:

$$\widehat{K} = \frac{p_0 - p_e}{1 - p_e}$$

where

p_0 = Accuracy A of observed agreement, $\frac{\sum X_{ii}}{N}$

p_e = estimate of chance agreement, $\frac{\sum X_{i+} X_{+i}}{N^2}$

Model (Random Forest) uncertainty

Shannon Entropy (Shannon, 2001; Shadman Roodposhti et al., 2019)

$$e_x = -\sum_{i=1}^{h} P_i log_2 P_i$$

where P_i, is the probability of class membership for class labels h. Note that the choice of the logarithm base is irrelevant as it only influences e_x units.

Supervised classification

Machine Learning Classification: brief

With regard to supervised learning, if otherwise unspecified, the procedures herein reported and briefly described are taken from the book *"An introduction to statistical learning"* by James et al. (2013).

Decision tree

A decision tree is the elemental unit of Machine Learning classification. The decision tree can be used for classification (using categorical themes, factors) and regression (numerical, continuous data). Its structure represents observations of a given item (branches; ramification of predictors that lead to a decision) and conclusions of the item's target class, or value (leaves; class labels). The problem that the decision tree seeks to solve is how a target variable Y can be predicted and generalised based on a combination of x predictor features.

Bagging

Bootstrap and aggregating, collectively referred to as *Bagging* (Breiman, 2001), is a machine learning aggregation meta-algorithm. It was designed to improve stability and accuracy by reducing variance and bypassing overfitting of tree-based statistical classifiers. The technique enhances the classification performance by aggregating the randomly derived predictions of several uncorrelated *weak learners* (i.e. one decision tree).

Bootstrap sample

Considering a number of data points, from the training set D of size n, m new training datasets D_i of size n_i are generated by sampling from D with replacement. Sampling with replacement enables to select independent samples, each having the same chance of being selected.

Aggregation

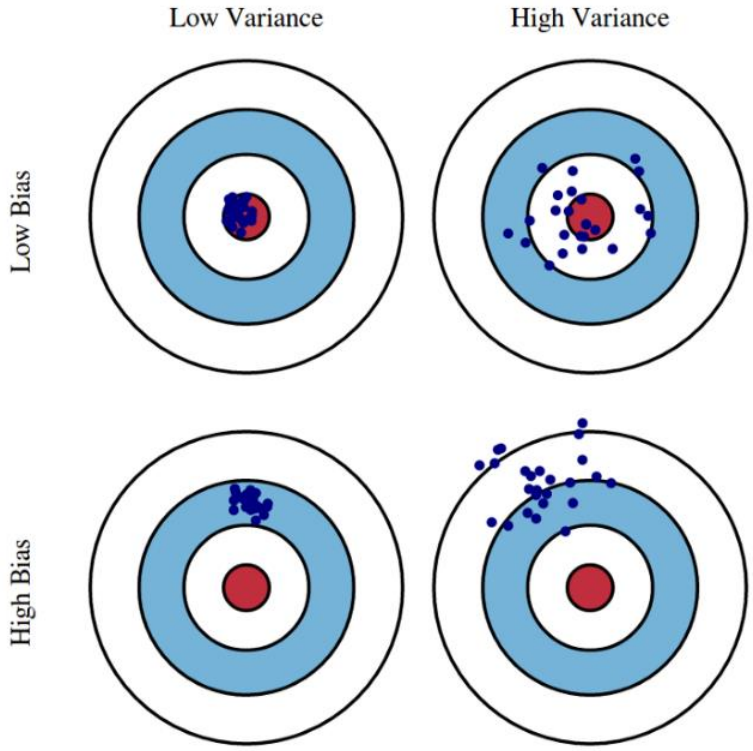
In turn, m models (for example a number of decision trees in a Random Forest classification) are fitted with m bootstrapped samples and the votes aggregated by majority voting. Considering a training set $X = x_1, \dots, x_n$ with associated predictors $Y = y_1, \dots, y_n$ a $for\ loop$ will B times take a bootstrapped subsample off the training dataset and fit the tree so that:

$$for\ b = 1, \dots, B$$

> Sample with replacement n observations from X and Y i.e. X_b and Y_b and train the classification tree f_b.

Random Forest

Registered by Breiman (2001), Random Forest is a trade mark of an ensemble Machine Learning algorithm used for classification and/or regression in various disciplines ranging from stock market to seafloor classification. It combines the decisions of several classification decision trees (constructed through the bagging process previously described). Random Forest is designed to bypass the inherent tendency of decision trees to overfitting the observations of the training set. In addition to the random bootstrap sampling of the training samples (leaving part of them Out of Bag OOB for internal cross-validation and variable importance estimates), Random Forest takes a bootstrapped sample of the variables used to construct the decisions at the nodes of each tree (ramification). Thus, there are two elements of randomness in the construction process of each tree in the forest (at both rows and columns levels). The extra bagging process results in a decorrelating effect, so that each tree will be constructed on different input information, reducing bias in the selection of training samples and stabilising the prediction (i.e. reducing the variance).



Schematic conceptualisation of bias and variance in the predictive performance of a classification algorithm.

Cross validation

Also referred to as OOB test (Out Of Bag), cross validation is a model validation approach to understand how well a given model is able (i.e. accurate) to generalise the predictions to an independent dataset. The target of cross validation is to test the accuracy of the model to predict unseen data on different shuffled sets of input training

and validation points. Here, it is used in supervised prediction using the Random Forest to test the predictive performance of the approach based on different combinations of hyperparameters i.e. $mtry$, the number of bagged features and $ntree$, the number of trees grown in the forest.

k-fold cross-validation

A form of class-validation which begins with a random partition of the training set into k evenly distributed subsamples. One of these is kept apart for validation (or testing) and the rest $k-1$ are used for training the model. $k \cdot fold$ (in this case $k = 10$) then repeats k times the process, validating the outcome on each of the k subsets. Accuracy estimates are then averaged to obtain the final cross-validated score. As an example, considering a cross-validation with $k = 2$; randomly shuffle (permute) the overall dataset D into subsets d_0 and d_1, and train on d_0 and cross validate on d_1 and vice versa.

Unsupervised classification

Unsupervised learning seeks to maximise within cluster homogeneity and between cluster dissimilarity without any *a priori* information of the underlying classes. The fundamental parameters that need specifying are; (1) the "seed" i.e. random initialisation partition number for the initial allocation of proto-cluster centroids; (2) the number of clusters, k, and (3) the number of initial proto-cluster centroid allocation from which the iterative clustering procedure can start (e.g. $nstart$ in R native k-means algorithm – Hartigan and Wong (1979)).

A "seed" is a paramount feature in unsupervised clustering (or any random operation) as it allows results to be reproduced exactly, starting from the same set of initially allocated points. The seed is the number from which a sequence of random numbers (e.g. in k-means these are the cluster centroids or proto-clusters) are generated.

Distance

A distance (similarity) function is required to quantify the affinity between two data points i.e. the length of the segment separating two points, p and q; $\overline{(pq)}$.

Euclidean distance

The Euclidean distance quantifies distances between two points in a n-dimensional space (either univariate or multivariate). The Euclidean distance between two backscatter pixels p_i and q_i is computed using the Pythagorean formula:

$$D(p_{i,}q_i) = \sqrt{\sum_{i=1}^{n}(q_i - p_i)^2}$$

k-means

K-means is a partitive clustering algorithm that seeks to allocate each n datapoint of each m feature vector into k clusters $S_i(i =, \dots, k)$. Iteratively, the process minimises

the sum of squared Euclidean distances between the datapoints and the overall average within the cluster (the proto-cluster or centroid). Thus, the minimisation problem reads:

$$min \sum_{i=1}^{k} \sum_{x_s \in S_i} \|X_s - c_i\|^2$$

Where X_s is a datapoint in cluster i, and c_i is the centroid of cluster i.

Clustering-aids to find k

Amongst the most important aspects in unsupervised learning, is that of objectively (statistically) finding the optimal number of clusters i.e. the number of clusters that the data can support (natural groupings) and that the algorithm can identify without creating artificial clusters. This exercise also confirms the number of classes which have been expertly categorised.

Within Group Sum of Squared Distances (Sum of Squared Error) – Elbow

This technique (and the one presented subsequently) is a wrapper function to the k-means algorithm and seeks to quantify the clustering optimum by considering a metric expressing the overall distance of a set of data points from their cluster. The metric is computed for a set of k-means solutions i.e. varying the number of clusters from two to ten. The clustering solution after which the metric does not improve is chosen as the optimum value of k.

For each point i take the squared distances from the points of the nearest cluster and sum them so that

$$SSE = \sum_{i=1}^{K} \sum_{x \in C_i} dist^2(m_i, x)$$

where x is a point in cluster C_i and m_i the centroid point in the cluster.

Silhouette coefficient (Rousseeuw, 1987)

Similarly to the previous clustering-aid, though evaluating the similarity of each point in their cluster with the similarity of that point to other (neighbouring) clusters.

First, a coefficient quantifying the “goodness of fit” (average Euclidean distance between i ($i \in C_i$) and points in its cluster) of each point i to its own cluster C_i (noted $a(i)$) is computed

$$a(i) = \frac{1}{\|C_i\| - 1} \sum_{j \in C_i, i \neq j} d(i,j)$$

where $d(i,j)$ is the distance between points i and j in cluster C_i

Secondly, define the mean distance of point i to another k cluster, defined as the mean distance of point i to the points of another k cluster c (noted $b(i)$). This is defined as the smallest possible mean distance of point i to the overall points of each other cluster. Therefore, a neighbouring cluster can be assigned to point i.

$$b(i) = \min_{i \neq j} \frac{1}{\|C_i\|} \sum_{j \in C_j} d(i,j)$$

These values can be used to form a *silhouette* coefficient $s(i)$ of point for point i

$$s(i) = \begin{cases} 1 - \frac{a(i)}{b(i)}, & if a(i) < b(i) \\ 0, & if\ a(i) = b(i) \\ \frac{b(i)}{a(i)} - 1, & if\ a(i) > b(i) \end{cases}$$

So that,

$$-1 \leq s(i) \leq 1$$

$a(i)$ must be $<<\ b(i)$ for $s(i)$ to approach 1.

Goodness of fit testing (fitting)

This approach is after Simons and Snellen (2009) though applied to the PDF of the backscatter histogram at 45°.

Reduced $\boldsymbol{\chi^2}$

Defined as χ^2 by degrees of freedom,

$$\chi^2_v = \frac{\chi^2}{v},$$

Where v are the degrees of freedom, $v = n - m$, n observations minus the number of fitted parameters m (3 per Gaussian considering Chapter 6) and where χ^2 , is the weighted sum of squared deviations,

$$\chi^2 = \sum_i \frac{(O_i - C_i)^2}{\sigma_i^2}$$

where σ_i^2 is the variance and O_i and C_i are the observed and modelled data respectively.

Fitting procedure

The method is implemented after Simons and Snellen (2009) though the approach used in this thesis does not refer to any physical prior and should be only considered as a statistical fitting procedure to find k.

Assuming that the Gaussian PDF approximates m seafloor type from the PDF of the backscatter data at $45°$, the fitting to the backscatter PDF refers to the sum of m Gaussians PDFs,

$$f(y_i|X) = \sum_{k=1}^{m} C_k \exp\left(-\frac{(y_j - \bar{y}_k)^2}{2\sigma_{y_k}^2}\right)$$

where, $f(y_i|X)$ is the model value at backscatter value y_j and the vector x contains the unknow parameters, $x = (\bar{y}_1, \dots, \bar{y}_m, \sigma_{y1}, \dots, \sigma_{ym}, C_i, \dots, C_m)^T$.

The number of m Gaussians is unknown (it is the target of the exercise) and it is therefore a free parameter. The unknown parameters for each gaussian (in vector x) are identified by, in a least square sense, maximise the fit between model and data:

$$\underbrace{min}_{x} \sum_{j=1}^{M} (p_j - f(y_i|x))^2$$

where, M is the number of bins and the search range for the Standard Deviation is herein arbitrarily set between 0 and 5 and p_j is the PDF at bin j (since the histogram is normalised). In turn, m is to be found by increasing number of m fitted against the goodness of fit. Where increasing m does not lead to an improvement of the fit, the number of classes is established.

Morphometric and textural MBES derivatives in a neighbourhood

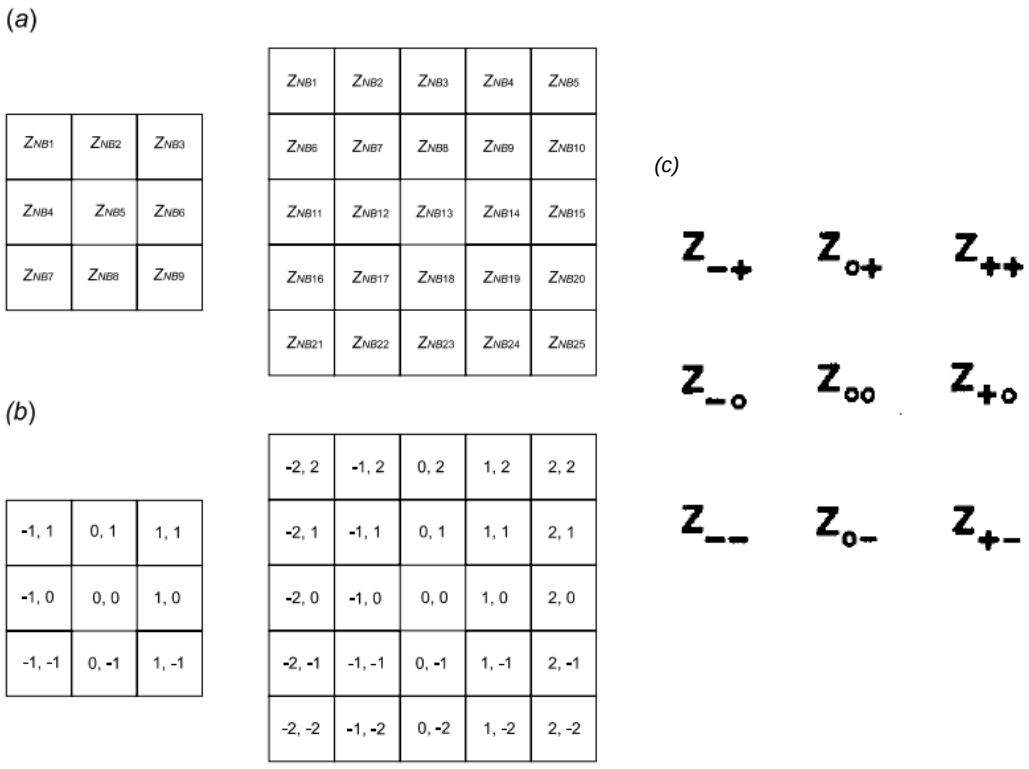
These variables where computed using the raster (Hijmans et al., 2014 and references therein) and GLCM (Zvoleff, 2016 and references therein) R packages.

Geomorphometry – Bathymetric grids

Definition: Science of quantitative land surface analysis (Pike et al., 2009).

All calculations are based on DTMs in raster 32-bit float format.

After Pike et al. (2009; a and b) and Horn (1981; c), the representation of a neighbourhood in a DTM and the notation for elevation of neighbouring points:



Common designation of neighbours in 3 x 3 or 5 x 5 windows (neighbours or kernels). In a) displaying unique identifiers; in b) by row and column separation of pixels from the central pixel and c) classical notation from pioneering studies (same as in b).

Slope

The slope gradient to the central pixel is derived as the average change in elevation. To achieve this, three steps are necessary (considering a $3\ x\ 3$ neighbour);

First, compute the difference in relative elevation in x and y directions, after which the gradient of the slope is given as the average of the two quadratic equations G and H,

$$G = \frac{Z_{NB3} + Z_{NB6} + Z_{NB9} - Z_{NB1} - Z_{NB4} - Z_{NB7}}{6\ \Delta s}$$

$$H = \frac{Z_{NB1} + Z_{NB2} + Z_{NB3} - Z_{NB7} - Z_{NB8} - Z_{NB9}}{6\,\Delta s}$$

With G being the first derivative in the x direction (df/dx) and H the first derivative in the y direction (df/dy). Z_{NB5} is the central pixel of the 3 x 3 neighbour for which the slope value is computed, Z_{NBi} are its neighbours $(n = 8)$ and Δs is the metric unit of the pixel size. At last, the slope as a tangent reads,

$$SLOPE = \sqrt{H^2 + G^2}$$

Topographic Position Index

The Topographic Position Index (TPI) is the terrestrial equivalent of the Bathymetric Position Index (BPI) used in marine Geomorphometry (Lundblad et al., 2006). The index quantifies the relative location of a feature in respect to the overall scene (i.e. digital surface), providing information of concavity, convexity and flatness. It is defined as:

$$BPI = < SCALEFACTOR > = int\left(Z_{grid} - focalmean\left(Z_{grid}, circle, r\right)\right) + 0.5)$$

Where the $SCALEFACTOR$ is the product of the size of the radius in metric (map) units and the DTM pixel resolution. Z_{grid} is the gridded surface and the $focalmean$ computes the average value of the grid cells within the specified radius r.

Terrain Ruggedness Index

The TRI index measures the local variability in seafloor terrain around a central pixel and is defined as (considering the equivalent notation in (b)).

$$TRI = \left([Z_{(-1,1)} - Z_{(-0,0)}] + [Z_{(0,1)} - Z_{(0,0)}] + [Z_{(1,1)} - Z_{(0,0)}] + [Z_{(-1,0)} - Z_{(0,0)}] + [Z_{(1,0)} - Z_{(0,0)}] + [Z_{(1,-1)} - Z_{(0,0)}] + [Z_{(0,-1)} - Z_{(0,0)}] + [Z_{(1,-1)} - Z_{(0,0)}]\right)/8$$

Rugosity

Another metric of terrain complexity derived by,

$$rugosity = \frac{3\ X\ 3\ surface\ area}{3\ X\ 3\ planar\ area}$$

Grey Level Co-occurrence Matrices (GLCMs) – Backscatter grids (in a window)

After Haralick and Shanmugam (1973) and Blondel and Sichi (2009).

Defined as the calculation of image textures.

Grey level co-occurrence matrices tabulate how often the various combinations of grey-levels, (NG_{0-255}), pixel values (say brightness, reflectance, reflectivity or else)

occur within a neighbourhood. Therefore, GLCM is a matrix tabulating the frequencies of neighbouring pairs of pixel values (dB) in a window.

The goal of GLCM analysis is to quantify (1) differences in NG, (2) define the area of change (window) and (3) the directionality of the event (whether omnidirectional or not). Computations are performed in shifts of $45°$ in the neighbourhood.

Tabulated frequencies of pairs of pixels (reference and neighbour pixels in a window) are then used to compute various image textures using different metrics (those reported hereafter are defined in Chapter 3, Table 3.2).

Entropy

$$Entropy = \sum_i \sum_j C_{ij}\, log\, C_{ij}$$

Contrast

$$Contrast = \sum_i \sum_j (i-j)^2\, C_{ij}$$

Dissimilarity

$$Dissimilarity = \sum_{i,j=0}^{N-1} c_{i,j}\, [i-j]$$

Variance

$$Variance = \sigma_i^2 \sum_{i,j=0}^{N-1} c_{i,j}\, [i-\mu_i]^2$$

Statistics – Backscatter (in a neighbourhood)

Mean

The sample mean (or average) is computed as the sum of all the points in the sample (the neighbourhood) divided by the total number of points in the sample,

$$\tilde{x} = \frac{1}{n}\sum_{i=1}^{n} x$$

where n is the sample size and x the values.

Minimum

The smallest value of a sorted vector of observations in a sample A (minimum entry)

$$\min A$$

Maximum

The largest value of a sorted vector of observations in a sample A (maximum entry)

$$\max A$$

Mode

The value occurring most frequently in the neighbourhood (most frequent entry)

Median

The median quantile in a distribution Q_2

Standard deviation

Measure of the spread of a set of points around their mean value, derived from the squared root of variance σ

$$\sigma = \frac{1}{n-1}\sum_{i=1}^{n}(x-\tilde{x})^2$$

and the standard deviation S

$$S = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(x-\tilde{x})^2}$$

Transmission loss (TL) in the active sonar equation

Two-way *Transmission Losses* (TL) including the term for spherical spreading and the absorption coefficient is reported hereafter

$$2TL = 40\ log\ R + \alpha R$$

Where R is the range, $40\ log$ is the spherical spreading to account for the decrease in acoustic intensity with distance from the source, the overall absorption coefficient α is itself the sum of three terms,

$$\alpha = \alpha_w + \alpha_{vs}$$

Where $\alpha_{vs} = \alpha_v + \alpha_s$

Absorption in seawater

After Francois and Garrison (1982a, b):

$$\alpha_w = A^1 P^1 \frac{f_1 f^2}{f_1^2 f^2} + A^1 P^2 \frac{f_2 f^2}{f_2^2 f^2} + A^3 P^3 f^2$$

where α is the attenuation expressed in dB/km, z is the depth in m; S is the salinity expressed in Practical Salinity Units (PSU), T is the temperature in Celsius degrees $C°$ and f is the frequency in kHz. A_i describes temperature dependencies, P_i the pressure dependencies f_i the relaxation frequencies of chemical reactions. Subscripts 1, 2 and 3 are the boric acid, magnesium sulphate and pure water absorption respectively reported hereafter:

Magnesium sulphate $M_g(SO_4)$

$$\begin{cases} A_2 = 21.44 \frac{S}{c} (1 + 0.025T) \\ P^2 = 1 - 1.37 \times 10^{-4} z + 6.2 \times 10^{-9} z^2 \\ f_2 = \frac{8.17 \times 10^{(8-1990)/(T+273))}}{1 + 0.0018(S-35)} \end{cases}$$

Boric acid $B(OH)_3$

$$\begin{cases} A^1 = \frac{8.86}{c} 10^{(0.78\, pH-5)} \\ P_1 = 1 \\ f_1 = 2.8 \sqrt{\frac{S}{35}} 10^{(4-\frac{1245}{T+273})} \\ c = 1412 + 3.21T + 1.19S + 0.0167z \end{cases}$$

Pure Water Viscosity

$$\begin{cases} P_3 = 1 - 3.83 \times 10^{-5} z + 4.9 \times 10^{-10} z^2 \\ T < 20°C \Rightarrow A_3 = 4.937 \times 10^{-4} - 2.59 \times 10^{-5} T + 9.11 \times 10^{-7} T^2 - 1.5 \times 10^{-8} T^3 \\ T > 20°C \Rightarrow A_3 = 3.964 \times 10^{-4} - 1.146 \times 10^{-5} T + 1.45 \times 10^{-7} T^2 - 6.5 \times 10^{-8} T^3 \end{cases}$$

Absorption due to suspended sediment

After Richards et al. (1996) and Hoitink and Hoekstra (2005) and assumptions and further references therein.

Alpha Viscosity

$$\alpha_v = \left(\frac{\epsilon k (\sigma - 1)^2}{2} \left[\frac{s}{s^2 + (\sigma + \delta)^2} \right] \right)$$

where $\delta = \frac{1}{2} \left[1 + \frac{9}{2\beta \langle \alpha_s \rangle} \right]$

and $s = \frac{9}{4\beta\langle\alpha_s\rangle}\left[1 + \frac{1}{\beta\langle\alpha_s\rangle}\right]$

$\beta = (\frac{\omega}{2v})^{1/2}$ is the oscillatory boundary layer depth, ϵ is the volume concentration of the Suspended Particulate Matter (SPM), k is the acoustic wave number ($^{2\pi}/_{\lambda}$), σ is the ratio of the densities of the sediment (ρ_s) and seawater (ρ_0) phases, $\langle\alpha_s\rangle$ is the mean radius of the particles, *v* is the kinematic viscosity of water and $\omega = 2\pi f$ where f is the operating frequency.

Alpha Scattering

$$\alpha_s = \left(\frac{\epsilon K_\alpha x^4}{\langle\alpha_s\rangle(1 + \xi x^2 + \frac{4}{3}K_\alpha x^4}\right)$$

where $K_\alpha = \frac{1}{6}\left(y_k^2 + \frac{\gamma_\rho^2}{3}\right)$, here $x = k\langle\alpha_s\rangle$ is the dimensionless parameter, y_k and y_p the compressibility and density contrasts between seawater and sediment. The remainder as previously.

References

A

Aldrich, R. C., 1975. Detecting disturbances in a forest environment. Photogramm. Eng. Remote Sensing, 41, 39–48.

Alevizos, E., Greinert, J., 2018. The Hyper-Angular Cube Concept for Improving the Spatial and Acoustic Resolution of MBES Backscatter Angular Response Analysis. Geosciences 8, 446.

Al-Habashneh, A.-A., Moloney, C., Gill, E., Huang, W., 2015. An Adaptive Method of Wave Spectrum Estimation Using X-Band Nautical Radar. Remote Sensing 7, 16537–16554.

Amiri-Simkooei, A., Snellen, M., Simons, D.G., 2012. Improving riverbed sediment classification using backscatter and depth residual features of multi-beam echo-sounder systems. J. Acoust. Soc. Am. 131, 3710–3725.

Anderson, J.T., Holliday, D.V., Kloser, R., Reid, D., Simrad, Y., Brown, C., Chapman, R., Coggan, R., Kieser, R., Michaels, L.W., Orlowski, A., Preston, J., Simmonds, J., Stepnowski, A. (Eds.), 2007. Acoustic Seabed Classification of Marine Physical and Biological Landscapes, ICES cooperative research report. Internat. Council for the Exploration of the Sea, Copenhagen.

Anderson, J.T., Van Holliday, D., Kloser, R., Reid, D.G., Simard, Y., 2008. Acoustic seabed classification: current practice and future directions. ICES J. Mar. Sci. 65, 1004–1011.

APL-UW., 1994. High-Frequency Ocean Environmental Acoustic Models Handbook (APL-UW TR 9407). Seattle, WA: Applied Physics Laboratory, University of Washington.

Arganda-Carreras, I., Kaynig, V., Rueden, C., Eliceiri, K.W., Schindelin, J., Cardona, A., Sebastian Seung, H., 2017. Trainable Weka Segmentation: a machine learning tool for microscopy pixel classification. Bioinformatics 33, 2424–2426.

Ashley, G.M., 1990. Classification of large-scale subaqueous bedforms; a new look at an old problem. J. Sediment. Res. 60, 160–172.

Assis, J., Narvaez, K., Haroun, R., 2007. Underwater towed video: a useful tool to rapidly assess elasmobranch populations in large marine protected areas. J. Coast. Conserv. 11, 153–157.

Augustin, J.-M., Lurton, X., 2005. Image amplitude calibration and processing for seafloor mapping sonars, in: Europe Oceans 2005. Presented at the Oceans 2005 - Europe, IEEE, Brest, France, 1, 698-701.

B

Baeye, M., Fettweis, M., 2015. In situ observations of suspended particulate matter plumes at an offshore wind farm, southern North Sea. Geo-Marine Letters 35, 247–255.

Baeye, M., Fettweis, M., Legrand, S., Dupont, Y., Van Lancker, V., 2012. Mine burial in the seabed of high-turbidity area — Findings of a first experiment. Cont. Shelf. Res. 43, 107–119.

Ban, Y., Yousif, O., 2016. Change Detection Techniques: A Review, in: Ban, Y. (Ed.), Multitemporal Remote Sensing. Springer International Publishing, Cham, 19–43.

Banko, G., 1998. A Review of Assessing the Accuracy of Classifications of Remotely Sensed Data and of Methods Including Remote Sensing Data in Forest Inventory. International Institute for Applied Systems Analysis. URL http://pure.iiasa.ac.at/5570/1/IR-98-081.pdf, (accessed 10.05.2018).

Barbier, E.B., Hacker, S.D., Kennedy, C., Koch, E.W., Stier, A.C., Silliman, B.R., 2011. The value of estuarine and coastal ecosystem services. Ecol. Monogr. 81, 169–193.

Bass, G.F., 1972. A history of seafaring based on underwater archaeology. Thames and Hudson, New York, USA.

Beaudoin J.D., Hughes Clarke J.E., van den Ameele E.J., Gardner J.V., 2002. Geometric and radiometric correction of multibeam backscatter derived from Reson 8101 systems. Canadian Hydrographic Conference 2002 Proceedings, Toronto, Ontario.

Beisiegel, K., Darr, A., Gogina, M., Zettler, M.L., 2017. Benefits and shortcomings of non-destructive benthic imagery for monitoring hard-bottom habitats. Mar. Pollut. Bull. 121, 5–15.

Belgian State., 2012. Determination of good environmental status and establishment of environmental targets for the Belgian marine waters. Federal Public Service Health Food Chain Safety and Environment. URL https://www.health.belgium.be/sites/default/files/uploads/fields/fpshealth_theme_file/19087665/Goede%20milieutoestand-MSFD%20EN.pdf, (accessed 04.05.2017).

Belgiu, M., Drăguţ, L., 2016. Random forest in remote sensing: A review of applications and future directions. ISPRS J. Photogramm. Remote Sens. 114, 24–31.

Bellec, V.K., Lancker, V.V., Degrendele, K., Roche, M., 2010. Geo-environmental Characterization of the Kwinte Bank. J. Coast. Res. 14.

Best, J., Simmons, S., Parsons, D., Oberg, K., Czuba, J., Malzone, C., 2010. A new methodology for the quantitative visualization of coherent flow structures in alluvial channels using multibeam echo-sounding (MBES). Geophys. Res. Let. 37.

Blomqvist, S., 1991. Quantitative sampling of soft-bottom sediments: problems and solutions. Mar. Ecol. Prog. Ser. 72, 295–304.

Blondel, P., 1996. Segmentation of the Mid-Atlantic Ridge south of the Azores, based on acoustic classification of TOBI data. Geol. Soc. Lond. Spec. Publ. 118, 17–28.

Blondel, P., Prampolini, M., Foglini, F., 2015. Acoustic textures and multibeam mapping of shallow marine habitats–examples from Eastern Malta. Proc. Inst. Acoust. 37, 250–257.

Blondel, P., Sichi, O.G., 2009. Textural analyses of multibeam sonar imagery from Stanton Banks, Northern Ireland continental shelf. Appl. Acoust. 70, 1288–1297.

Blott, S.J., Pye, K., 2001. GRADISTAT: a grain size distribution and statistics package for the analysis of unconsolidated sediments. Earth Surf. Process. Landf. 26, 1237–1248.

Boehme, H., Chotiros, N.P., Rolleigh, L.D., Pitt, S.P., Garcia, A.L., Goldsberry, T.G., Lamb, R.A., 1984. Acoustic backscattering at low grazing angles from the ocean bottom. I. Bottom backscattering strength. J. Acoust. Soc. Am. 75, S30–S30.

Boessenecker, R.W., Perry, F.A., Schmitt, J.G., 2014. Comparative taphonomy, taphofacies, and bonebeds of the Mio-Pliocene Purisima Formation, Central California: strong physical control on marine vertebrate preservation in shallow marine settings. PLoS One 9, e91419.

Borja, A., Elliott, M., Andersen, J.H., Cardoso, A.C., Carstensen, J., Ferreira, J.G., Heiskanen, A.-S., Marques, J.C., Neto, J.M., Teixeira, H., 2013. Good Environmental Status of marine ecosystems: what is it and how do we know when we have attained it? Mar. Pollut. Bull. 76, 16–27.

Boström, C., Pittman, S.J., Simenstad, C., Kneib, R.T., 2011. Seascape ecology of coastal biogenic habitats: advances, gaps, and challenges. Mar. Ecol. Prog. Ser. 427, 191–217.

Boyd, S.E., Coggan, R.A., Birchenough, S.N.R., Limpenny, D.S., Eastwood, P.E., Foster-Smith, R.L., Philpott, S., Meadows, W.J., James, J.W.C., Vanstaen, K., Soussi, S., Rogers, S., 2006. The role of seabed mapping techniques in environmental monitoring and management. Sci. Ser. Technical report, Cefas, 127-170.

Braimoh, A.K., 2006. Random and systematic land-cover transitions in northern Ghana. Agric. Ecosyst. Environ. 113, 254–263.

Breiman, L., 2001. Random forests. Mach. Learn. 45, 5–32.

Breine, N.T., De Backer, A., Van Colen, C., Moens, T., Hostens, K., Van Hoey, G., 2018. Structural and functional diversity of soft-bottom macrobenthic communities in the Southern North Sea. Estuar. Coast. Shelf Sci. 214, 173–184.

Breman, J., 2002. Marine geography: GIS for the oceans and seas. ESRI press, Inc. Redlands, California.

Briggs, K.B., Williams, K.L., Richardson, M.D., Jackson, D.R., 2001. Effects of Changing Roughness on Acoustic Scattering: (1) Natural Changes. Proc. Inst. Acoust. 23, 375–382.

Brown, C.J., Blondel, P., 2009. Developments in the application of multibeam sonar backscatter for seafloor habitat mapping. Appl. Acoust. 70, 1242–1247.

Brown, C.J., Sameoto, J.A., Smith, S.J., 2012. Multiple methods, maps, and management applications: Purpose made seafloor maps in support of ocean management. J. Sea Res. 72, 1–13.

Brown, C.J., Smith, S.J., Lawton, P., Anderson, J.T., 2011. Benthic habitat mapping: A review of progress towards improved understanding of the spatial ecology of the seafloor using acoustic techniques. Estuar. Coast. Shelf Sci. 92, 502–520.

Brown, R.J., Brisco, B., Ahern, F.J., Bjerkelund, C., Manore, M., Pultz, T.J., Singhroy, V., 1993. SAR application calibration requirements. Can. J. Remote Sens. 19, 193–203.

Buck, C.H., 2000. Alternative large-scale distributed targets for SAR elevation beam pattern characterization, in: Proc. ERS-Envisat Symposium. Gothenburg, Sweden, 16-20 october 2000, 1-7.

Buhl-Mortensen, L., Buhl-Mortensen, P., Dolan, M.F.J., Holte, B., 2015. The MAREANO programme – A full coverage mapping of the Norwegian off-shore benthic environment and fauna. Mar. Biol. Res. 11, 4–17.

Buscombe, D., Grams, P., 2018. Probabilistic Substrate Classification with Multispectral Acoustic Backscatter: A Comparison of Discriminative and Generative Models. Geosciences 8, 395.

Buscombe, D., Grams, P.E., Kaplinski, M.A., 2017. Compositional Signatures in Acoustic Backscatter Over Vegetated and Unvegetated Mixed Sand-Gravel Riverbeds: Acoustic river substrate classification. J. Geophys. Res. Earth Surf. 122, 1771–1793.

Butman, B., Bothner, M.H., Hathaway, J.C., Jenter, H.L., Knebel, H.J., Manheim, F.T., Signell, R.P., 1992. Contaminant transport and accumulation in Massachusetts Bay and Boston Harbour: a summary of US Geological Survey studies. US Geol. Surv. No. 92-202.

C

Callaway, R., Engelhard, G.H., Dann, J., Cotter, J., Rumohr, H., 2007. A century of North Sea epibenthos and trawling: comparison between 1902–1912, 1982–1985 and 2000. Mar. Ecol. Prog Ser. 346, 27–43.

Calvert, J., Strong, J.A., Service, M., McGonigle, C., Quinn, R., 2015. An evaluation of supervised and unsupervised classification techniques for marine benthic habitat mapping using multibeam echosounder data. ICES J. Mar. Sci. 72, 1498–1513.

Capolsini, P., Andréfouët, S., Rion, C., Payri, C., 2003. A comparison of Landsat ETM+, SPOT HRV, Ikonos, ASTER, and airborne MASTER data for coral reef habitat mapping in South Pacific islands. Can. J. Remote Sens. 29, 187-200.

Charette, M., Smith, W., 2010. The Volume of Earth's Ocean. Oceanography 23, 112–114.

Chen, Z., Hu, C., Muller-Karger, F.E., Luther, M.E., 2010. Short-term variability of suspended sediment and phytoplankton in Tampa Bay, Florida: observations

from a coastal oceanographic tower and ocean color satellites. Estuar. Coastal. Shelf. Sci 89, 62–72.

Chimienti, G., Angeletti, L., Rizzo, L., Tursi, A., Mastrototaro, F., 2018. ROV vs trawling approaches in the study of benthic communities: the case of *Pennatula rubra* (Cnidaria: Pennatulacea). J. Mar. Biol. Ass. 98, 1859–1869.

Coggan, R., Populus, J., White, J., Sheehan, K., Fitzpatrick, F., Piel, S., 2007. Review of standards and protocols for seabed habitat mapping. Mapp. Eur. Seabed Habitats MESH Peterb. UK.

Cohen, J., 1960. A coefficient of agreement for nominal scales. Educ. Psychol. Meas. 20, 37–46.

Colbo, K., Ross, T., Brown, C., Weber, T., 2014. A review of oceanographic applications of water column data from multibeam echosounders. Estuar. Coastal. Shelf. Sci. 145, 41–56.

Collier, J.S., Brown, C.J., 2005. Correlation of sides can backscatter with grain size distribution of surficial seabed sediments. Mar. Geol. 214, 431–449.

Congalton, R.G., 1991. A review of assessing the accuracy of classifications of remotely sensed data. Remote Sens. Environ. 37, 35–46.

Cook, E.J., Black, K.D., Sayer, M.D.J., Cromey, C.J., Angel, D.L., Spanier, E., Tsemel, A., Katz, T., Eden, N., Karakassis, I., 2006. The influence of caged mariculture on the early development of sublittoral fouling communities: a pan-European study. ICES J. Mar. Sci. 63, 637–649.

Coppin, P., Jonckheere, I., Nackaerts, K., Muys, B., Lambin, E., 2004. Review ArticleDigital change detection methods in ecosystem monitoring: a review. Int. J. Remote Sens. 25, 1565–1596.

Culloch, R., Bennet, F., Bald, J., Menchaca, I., Jessopp, M., Simas, T., 2015. Report on potential emerging innovative monitoring approaches, identifying potential reductions in monitoring costs and evaluation of existing long-term datasets. Deliverable 4.3. RICORE Project, 61.

D

Daniell, J., Siwabessy, J., Nichol, S., Brooke, B., 2015. Insights into environmental drivers of acoustic angular response using a self-organising map and hierarchical clustering. Geo-Mar. Lett. 35, 387–403.

Dannheim, J., Bergström, L., Birchenough, S.N.R., Brzana, R., Boon, A.R., Coolen, J.W.P., Dauvin, J.-C., De Mesel, I., Derweduwen, J., Gill, A.B., Hutchison, Z.L., Jackson, A.C., Janas, U., Martin, G., Raoux, A., Reubens, J., Rostin, L., Vanaverbeke, J., Wilding, T.A., Wilhelmsson, D., Degraer, S., 2019. Benthic effects of offshore renewables: identification of knowledge gaps and urgently needed research. ICES J. Mar. Sci.

Davies, C.E., Moss, D., Hill, M.O., 2004. EUNIS Habitat Classification Revised 2004. Report to EEA and European Topic Centre on Nature Protection and Biodiversity. October 2004. URL

http://eunis.eea.europa.eu/upload/EUNIS_2004_report.pdf, (accessed 18.03.2016)

Davoult, D., Dewarumez, J. M., Prygiel, J., Richard, A. 1988. Carte des peuplements benthiques de la partie française de la mer du Nord. IFREMER/Région Nord, Pas-de-Calais. 30 URL http://www.rebent.org/docs/metadata/05_HabitatsHistoriques/02_HabitatsLegendeAuteurs/01A_MerDuNord/ifr_peupl_Davoult_MerDuNord_1988_l2_p_EUNIS2004_Metadonnees.pdf, (accessed 25.05.2019).

De Backer, A., Hillewaert, H., Van Hoey, G., Wittoeck, J., Hostens, K. 2014. Structural and functional biological assessment of aggregate dredging intensity on the Belgian part of the North Sea, in: De Mol, L. et al. (Ed.) "Which future for the sand extraction in the Belgian part of the North Sea?". Study Day, 20 Oct. 2014, Belgium Pier - Blankenberge. URL https://economie.fgov.be/sites/default/files/Files/Entreprises/Sand/10-Articles-study-day_2014.pdf, (accessed 15.02.2019).

De Bisschop, J., 2016. Influence of water column properties on multibeam backscatter. (MSc), Ghent University.

de Campos Carvalho, R., de Oliveira Junior, A.M., Clarke, J.E.H., 2013. Proper environmental reduction for attenuation in multi-sector sonars, in: 2013 IEEE/OES Acoustics in Underwater Geosciences Symposium. Presented at the 2013 IEEE/OES Acoustics in Underwater Geosciences Symposium (RIO Acoustics), IEEE, Rio de Janeiro- RJ, Brazil, 1–6.

De Clercq, M., Chademenos, V., Van Lancker, V., Missiaen, T., 2016. A high-resolution DEM for the Top-Palaeogene surface of the Belgian Continental Shelf. J. Maps 12, 1047–1054.

De Clercq, M., Missiaen, T., Wallinga, J., Zurita Hurtado, O., Versendaal, A., Mathys, M., De Batist, M., 2018. A well-preserved Eemian incised-valley fill in the southern North Sea Basin, Belgian Continental Shelf - Coastal Plain: Implications for northwest European landscape evolution: A Belgian Eemian valley fill with European landscape implications. Earth Surf. Process. Landf. 43, 1913–1942.

De Falco, G., Tonielli, R., Di Martino, G., Innangi, S., Simeone, S., Parnum, I.M., 2010. Relationships between multibeam backscatter, sediment grain size and Posidonia oceanica seagrass distribution. Cont. Shelf Res. 30, 1941–1950.

De Mesel, I., Van Lancker, V., Kapasakali, D., **Montereale-Gavazzi**, G., Kerckhof, F., Van Hoey, G., Wittoeck, J., Hillewaert, H., Ranson, J., Vanelslander, B., Hostens, K. 2017. Analysis of the current status of the benthic substrate fauna in the soil protection zones in the Flemish Banks. Operational Directorate Natural Environment, OD Nature / Institute for Agricultural and Fisheries Research: Brussels, Ostend. URL http://pure.ilvo.vlaanderen.be/portal/files/5906504/T0_Zone1_Zone3_RBINS_ILVO.pdf, (accessed 10.02.2019).

De Moustier, C., 1986. Beyond bathymetry: Mapping acoustic backscattering from the deep seafloor with Sea Beam. J. Acoust. Soc. Am. 79, 316–331.

De Moustier, C., Alexandrou, D., 1991. Angular dependence of 12-kHz seafloor acoustic backscatter. J. Acoust. Soc. Am. 90, 522–531.

De Moustier, C., Matsumoto, H., 1993. Seafloor acoustic remote sensing with multibeam echo-sounders and bathymetric sides can sonar systems. Mar. Geophys. Res. 15, 27–42.

De, C., Chakraborty, B., 2011. Model-based acoustic remote sensing of seafloor characteristics. IEEE Trans. Geosci. Remote Sens. 49, 3868–3877.

Degraer, S., 1999. Macrobenthos of shallow marine habitats (Belgian coast) and its use in coastal zone management. (PhD), Ghent University.

Degraer, S., Brabant, R., Rumes, B., 2013. Environmental impacts of offshore wind farms in the Belgian part of the North Sea: Learning from the past to optimise future monitoring programmes. Royal Belgian Institute of Natural Sciences (RBINS), Operational Directorate Natural Environment. Marine Ecology and Management Section: Brussels. URL https://tethys.pnnl.gov/publications/environmental-impacts-offshore-wind-farms-belgian-part-north-sea-learning-past-optimise, (accessed 02.02.2019).

Degraer, S., Brabant, R., Rumes, B., Vigin, L., (Ed.) 2016. Environmental impacts of offshore wind farms in the Belgian part of the North Sea: Environmental impact monitoring reloaded. Royal Belgian Institute of Natural Sciences, OD Natural Environment, Marine Ecology and Management Section: Brussels. ISBN 978-90-8264-120-2. IX, 287.

Degraer, S., Moerkerke, G., Rabaut, M., Van Hoey, G., Du Four, I., Vincx, M., Henriet, J.-P., Van Lancker, V., 2008. Very-high resolution side-scan sonar mapping of biogenic reefs of the tube-worm Lanice conchilega. Remote Sens. Environ. 112, 3323–3328.

Degraer, S., Verfaillie, E., Willems, W., Adriaens, E., Vincx, M., Van Lancker, V., 2008. Habitat suitability modelling as a mapping tool for macrobenthic communities: An example from the Belgian part of the North Sea. Cont. Shelf. Res. 28, 369–379.

Deleu, S., Van Lancker, V., 2007. Geological setting of gravel occurrences on the Belgian part of the North Sea. Annex Van Lancker Al2007 Manag. Res. Budgeting Aggreg. Shelf Seas Relat. End-Users Marebasse Final Sci. Rep. Belg. Sci. Policy 101–115.

Diaz, R.J., Solan, M., Valente, R.M., 2004. A review of approaches for classifying benthic habitats and evaluating habitat quality. J. Environ. Econ. Manag. 73, 165–181.

Dierssen, H.M., Theberge Jr, A.E., 2014. Bathymetry: History of seafloor mapping. Encycl. Ocean Sci. 1-5

Diesing, M., Green, S.L., Stephens, D., Lark, R.M., Stewart, H.A., Dove, D., 2014. Mapping seabed sediments: comparison of manual, geostatistical, object-based image analysis and machine learning approaches. Cont. Shelf Res. 84, 107–119.

Diesing, M., Stephens, D., 2015. A multi-model ensemble approach to seabed mapping. J. Sea Res. 100, 62–69.

Douvere, F., 2008. The importance of marine spatial planning in advancing ecosystem-based sea use management. Mar. Policy 32, 762–771.

Douvere, F., Maes, F., Vanhulle, A., Schrijvers, J., 2007. The role of marine spatial planning in sea use management: The Belgian case. Mar. Pol. 31, 182–191.

Du Four, I., Van Lancker, V., 2008. Changes of sedimentological patterns and morphological features due to the disposal of dredge spoil and the regeneration after cessation of the disposal activities. Mar. Geol. 255, 15–29.

E

EC/2008/56. 2008. Marine Strategy Framework Directive 2008/56/EC of the European Parliament and of the Council of 17 June 2008 establishing a framework for community action in the field of marine environmental policy (Text with EEA relevance). URL http://eurlex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32008L0056&qid=1492697872524&from=en, (accessed 04.05.2017).

Eleftherakis, D., 2013. Classifying sediments on Dutch riverbeds using multi-beam echo-sounder systems. (PhD), Technical University Delft.

Eleftherakis, D., Berger, L., Le Bouffant, N., Pacault, A., Augustin, J.-M., Lurton, X., 2018. Backscatter calibration of high-frequency multibeam echosounder using a reference single-beam system, on natural seafloor. Mar. Geophys. Res. 39, 55–73.

Ernstsen, V.B., Noormets, R., Hebbeln, D., Bartholomä, A., Flemming, B.W., 2006a. Precision of high-resolution multibeam echo sounding coupled with high-accuracy positioning in a shallow water coastal environment. Geo-Mar. Lett. 26, 141–149.

Ernstsen, V.B., Noormets, R., Winter, C., Hebbeln, D., Bartholomä, A., Flemming, B.W., Bartholdy, J., 2006b. Quantification of dune dynamics during a tidal cycle in an inlet channel of the Danish Wadden Sea. Geo-Mar. Lett. 26, 151–163.

Evans, D. et alias. 2016 (revised 2017). Revising the marine section of the EUNIS Habitat classification - Report of a workshop held at the European Topic Centre on Biological Diversity, 12 and 13 May 2016. URL ETC/BD Working Paper Nº A/2016, revised 2017, (accessed 30.05.2019).

F

Feldens, P., Schulze, I., Papenmeier, S., Schönke, M., Schneider von Deimling, J., 2018. Improved Interpretation of Marine Sedimentary Environments Using Multi-Frequency Multibeam Backscatter Data. Geosciences 8, 214.

Ferrini, V.L., Flood, R.D., 2006. The effects of fine-scale surface roughness and grain size on 300 kHz multibeam backscatter intensity in sandy marine sedimentary environments. Mar. Geol. 228, 153–172.

Fettweis, M., Baeye, M., 2015. Seasonal variation in concentration, size, and settling velocity of muddy marine flocs in the benthic boundary layer: Seasonality of SPM concentration. J. Geophys. Res. Oceans 120, 5648–5667.

Fettweis, M., Baeye, M., Lee, B.J., Chen, P., Yu, J.C.S., 2012. Hydro-meteorological influences and multimodal suspended particle size distributions in the Belgian nearshore area (southern North Sea). Geo-Mar. Lett. 32, 123–137.

Fettweis, M., Francken, F., Pison, V., Van den Eynde, D., 2006. Suspended particulate matter dynamics and aggregate sizes in a high turbidity area. Mar. Geol. 235, 63–74.

Fettweis, M., Houziaux, J.-S., Du Four, I., Van Lancker, V., Baeteman, C., Mathys, M., Van den Eynde, D., Francken, F., Wartel, S., 2009. Long-term influence of maritime access works on the distribution of cohesive sediments: analysis of historical and recent data from the Belgian nearshore area (southern North Sea). Geo-Marine Letters 29, 321–330.

Fettweis, M., Lee, B., 2017. Spatial and Seasonal Variation of Biomineral Suspended Particulate Matter Properties in High-Turbid Nearshore and Low-Turbid Offshore Zones. Water 9, 694.

Fezzani, R., Berger, L., 2018. Analysis of calibrated seafloor backscatter for habitat classification methodology and case study of 158 spots in the Bay of Biscay and Celtic Sea. Mar. Geophys. Res. 39, 169–181.

Fish, J.P., Carr, H.A., 1990. Sound underwater images: a guide to the generation and interpretation of side scan sonar data. Lower Cape Publishing Company.

Floricioiu, D., Rott, H., 2001. Seasonal and short-term variability of multifrequency, polarimetric radar backscatter of Alpine terrain from SIR-C/X-SAR and AIRSAR data. IEEE Trans. Geosci. Remote Sens. 39, 2634–2648.

Fogarin, S., Madricardo, F., Zaggia, L., **Montereale-Gavazzi**, G., Sigovini, M., Kruss, A., Lorenzetti, G., Manfe, G., Petrizzo, A., Molinaroli, E., Trincardi, F., 2019. Tidal Inlets in the Anthropocene: geomorphology and benthic habitats of the Chioggia inlet, Venice Lagoon (Italy). Earth Surf. Process. Landf. (in press).

Folk, R.L., 1954. The Distinction between Grain Size and Mineral Composition in Sedimentary-Rock Nomenclature. J. Geol. 62, 344–359.

Fonseca, L., Calder, B., 2005. Geocoder: an efficient backscatter map constructor. Centre for Coastal and ocean Mapping, University of New Hampshire. URL http://ushydro.thsoa.org/hy05/08_3.pdf, (accessed 02.02.2019).

Fonseca, L., E., Calder, B., R., 2007. Clustering Acoustic Backscatter in the Angular Response Space. U.S. Hydrographic Conference. URL https://scholars.unh.edu/ccom/384, (accessed 02.02.2019).

Fonseca, L., Lurton, W., Fezzani, R., Augustin, J-M., Berger, L. 2019. Some Practical recommendations for averaging acoustic backscatter strength. Geological and Biological Marine Habitat Mapping Conference, International Symposium, St. Petersburg, Russia, 13-17 May 2019.

Fonseca, L., Mayer, L., 2007. Remote estimation of surficial seafloor properties through the application Angular Range Analysis to multibeam sonar data. Mar. Geophys. Res. 28, 119–126.

Fonseca, L., Mayer, L., Orange, D., Driscoll, N., 2002. The high-frequency backscattering angular response of gassy sediments: model/data comparison from the Eel River Margin, California. J. Acoust. Soc. Am. 111, 2621–2631.

Foody, G.M., 2002. Status of land cover classification accuracy assessment. Remote Sens. Environ. 80, 185–201.

Foody, G.M., 2004. Thematic map comparison. Photogramm. Eng. Remote Sens. 70, 627–633.

Foody, G.M., Boyd, D.S., 1999. Detection of partial land cover change associated with the migration of inter-class transitional zones. International J. Remote Sens. 20, 2723–2740.

Foody, G.M., Boyd, D.S., Sanchez-Hernandez, C., 2007. Mapping a specific class with an ensemble of classifiers. Int. J. Remote Sens. 28, 1733–1746.

Francois, R.E., Garrison, G.R., 1982a. Sound absorption based on ocean measurements: Part I: Pure water and magnesium sulphate contributions. J. Acoust. Soc. Am. 72, 896–907.

Francois, R.E., Garrison, G.R., 1982b. Sound absorption based on ocean measurements. Part II: Boric acid contribution and equation for total absorption. J. Acoust. Soc. Am. 72, 1879–1890.

Frost, G.L., 2001. Inventing schemes and strategies: The making and selling of the Fessenden oscillator. Technol. Cult. 42, 462–488.

G

Gaida, T., Tengku Ali, T., Snellen, M., Amiri-Simkooei, A., van Dijk, T., Simons, D., 2018. A Multispectral Bayesian Classification Method for Increased Acoustic Discrimination of Seabed Sediments Using Multi-Frequency Multibeam Backscatter Data. Geosciences 8, 455.

Galparsoro, I., Borja, Á., Bald, J., Liria, P., Chust, G., 2009. Predicting suitable habitat for the European lobster (Homarus gammarus), on the Basque continental shelf (Bay of Biscay), using Ecological-Niche Factor Analysis. Ecol. Modell. 220, 556–567.

Galparsoro, I., Connor, D.W., Borja, Á., Aish, A., Amorim, P., Bajjouk, T., Chambers, C., Coggan, R., Dirberg, G., Ellwood, H., Evans, D., Goodin, K.L., Grehan, A., Haldin, J., Howell, K., Jenkins, C., Michez, N., Mo, G., Buhl-Mortensen, P., Pearce, B., Populus, J., Salomidi, M., Sánchez, F., Serrano, A., Shumchenia, E., Tempera, F., Vasquez, M., 2012. Using EUNIS habitat classification for benthic mapping in European seas: Present concerns and future needs. Mar. Pollut. Bull. 64, 2630–2638.

Gavrilov, A.N., Siwabessy, P.J.W., Parnum, I.M., 2005. Multibeam echo sounder backscatter analysis, Centre for Marine Science and Technology, Perth, Australia, CA3.03.

Gazis, I.-Z., Schoening, T., Alevizos, E., Greinert, J., 2018. Quantitative mapping and predictive modeling of Mn nodules' distribution from hydroacoustic and optical

AUV data linked by random forests machine learning. Biogeosciences 15, 7347–7377.

Gilkinson, K.D., Fader, G.B.J., Gordon Jr, D.C., Charron, R., McKeown, D., Roddick, D., Kenchington, E.L.R., MacIsaac, K., Bourbonnais, C., Vass, P., 2003. Immediate and longer-term impacts of hydraulic clam dredging on an offshore sandy seabed: effects on physical habitat and processes of recovery. Cont. Shelf. Res. 23, 1315–1336.

Gille, B., 1966. Engineers of the Renaissance. (London: Lund Humphries), translated from the French *Les Ingenieurs de la Renaissance* (1964, Paris: Hermann).

Gilson, G., 1907. Exploration de la Mer sur les côtes de Belgique. Première série. Recherches sur le milieu marin et ses variations au voisinage de la côte Belge. In: Mémoires du Musée Royal d'Histoire Naturelle de Belgique, IV (distribue en 1911), Bruxelles, 81.

Goff, J.A., Kraft, B.J., Mayer, L.A., Schock, S.G., Sommerfield, C.K., Olson, H.C., Gulick, S.P., Nordfjord, S., 2004. Seabed characterization on the New Jersey middle and outer shelf: correlatability and spatial variability of seafloor sediment properties. Mar. Geol. 209, 147–172.

Goff, J.A., Olson, H.C., Duncan, C.S., 2000. Correlation of side-scan backscatter intensity with grain-size distribution of shelf sediments, New Jersey margin. Geo-Marine Letters 20, 43–49.

Goldsmith, F.B., 2012. Monitoring for conservation and ecology. Springer Science & Business Media.

Gong, P., Wang, J., Yu, L., Zhao, Yongchao, Zhao, Yuanyuan, Liang, L., Niu, Z., Huang, X., Fu, H., Liu, S., 2013. Finer resolution observation and monitoring of global land cover: First mapping results with Landsat TM and ETM+ data. International J. Remote Sens. 34, 2607–2654.

Gonima, L., 1993. Simple algorithm for the atmospheric correction of reflectance images. Int. J. Remote Sens. 14, 1179–1187.

Gorska, N., Kowalska-Duda, E., Pniewski, F., Latała, A., 2018. On diel variability of marine sediment backscattering properties caused by microphytobenthos photosynthesis: Impact of environmental factors. J. Mar. Syst. 182, 1–11.

Gowdy, J., Mesner, S., 1998. The evolution of Georgescu-Roegen's bioeconomics. Rev. Soc. Econ. 56, 136–156.

Greene, H., O'Connell, V., Brylinsky, C., Reynolds, J., 2008. Marine Benthic Habitat Classification: What's Best for Alaska? in: Reynolds, J., Greene, H. (Eds.), Marine Habitat Mapping Technology for Alaska. Alaska Sea Grant, University of Alaska Fairbanks, 169–184.

Guelorget, O., Perthuisot, J.P., 1992. Paralic ecosystems. Biological organization and functioning. *Vie et Milieu* 42, 215–251.

Gueriot, D., Chedru, J., Daniel, S., Maillard, E., 2000. The patch test: a comprehensive calibration tool for multibeam echosounders, in: OCEANS 2000 MTS/IEEE Conference and Exhibition. Conference Proceedings (Cat. No. 00CH37158). IEEE, 1655–1661.

Guinan, J., Brown, C., Dolan, M.F., Grehan, A.J., 2009. Ecological niche modelling of the distribution of cold-water coral habitat using underwater remote sensing data. Ecol. Inform. 4, 83–92.

Guisan, A., Tingley, R., Baumgartner, J.B., Naujokaitis-Lewis, I., Sutcliffe, P.R., Tulloch, A.I., Regan, T.J., Brotons, L., McDonald-Madden, E., Mantyka-Pringle, C., 2013. Predicting species distributions for conservation decisions. Ecol. Lett. 16, 1424–1435.

H

Hademenos, V., Stafleu, J., Missiaen, T., Kint, L., Van Lancker, V.R.M., 2019. 3D subsurface characterisation of the Belgian Continental Shelf: a new voxel modelling approach. Neth. J. Geosci., 1–17.

Hall, L.S., Krausman, P.R., Morrison, M.L., 1997. The habitat concept and a plea for standard terminology. Wildl. Soc. Bull., 173–182.

Halpern, B.S., Frazier, M., Potapenko, J., Casey, K.S., Koenig, K., Longo, C., Lowndes, J.S., Rockwood, R.C., Selig, E.R., Selkoe, K.A., Walbridge, S., 2015. Spatial and temporal changes in cumulative human impacts on the world's ocean. Nat. Commun. 6, 7615.

Halpern, B.S., Walbridge, S., Selkoe, K.A., Kappel, C.V., Micheli, F., D'Agrosa, C., Bruno, J.F., Casey, K.S., Ebert, C., Fox, H.E., Fujita, R., Heinemann, D., Lenihan, H.S., Madin, E.M.P., Perry, M.T., Selig, E.R., Spalding, M., Steneck, R., Watson, R., 2008. A Global Map of Human Impact on Marine Ecosystems. Science 319, 948–952.

Hamilton, E.L., 1980. Geoacoustic modeling of the sea floor. J. Acoust. Soc. Am. 68, 1313–1340.

Hammerstad, E., 2000. EM Technical Note: Backscattering and Seabed Image Reflectivity. Horten, Norway: Kongsberg Maritime AS. Technical note. URL https://www.km.kongsberg.com/ks/web/nokbg0397.nsf/AllWeb/C2AE0703809C1FA5C1257B580044DD83/$file/EM_technical_note_web_BackscatteringSeabedImageReflectivity.pdf, (accessed 01.03.2018)

Hansen, M.C., Potapov, P.V., Moore, R., Hancher, M., Turubanova, S.A.A., Tyukavina, A., Thau, D., Stehman, S.V., Goetz, S.J., Loveland, T.R., 2013. High-resolution global maps of 21st-century forest cover change. Science 342, 850–853.

Hanson, H., Brampton, A., Capobianco, M., Dette, H.H., Hamm, L., Laustrup, C., Lechuga, A., Spanhoff, R., 2002. Beach nourishment projects, practices, and objectives—a European overview. Coast. Eng. 47, 81–111.

Haralick, R.M., Shanmugam, K., 1973. Textural features for image classification. IEEE Trans. Syst. Man Cybern., 610–621.

Harris, P.T., 2012. Surrogacy, in: Seafloor Geomorphology as Benthic Habitat, 1st ed. Elsevier.

Hartigan, J.A., Wong, M.A., 1979. Algorithm AS 136: A K-Means Clustering Algorithm. Appl. Stat. 28, 100.

Hasan, R., Ierodiaconou, D., Laurenson, L., Schimel, A., 2014. Integrating Multibeam Backscatter Angular Response, Mosaic and Bathymetry Data for Benthic Habitat Mapping. PLoS ONE 9, e97339.

Hasan, R., Mohd Razali, M., Shamsudin, S., A. 2016. Sediment classification from multibeam backscatter images using simple histogram analysis. HYDRO 2016, Rostock Warnemundie, 08-10 November 2016.

Hasan, R.C., Ierodiaconou, D., Laurenson, L., Schimel, A., 2014. Integrating multibeam backscatter angular response, mosaic and bathymetry data for benthic habitat mapping. Plos One 9, e97339.

Hasan, R.C., Ierodiaconou, D., Monk, J., 2012. Evaluation of four supervised learning methods for benthic habitat mapping using backscatter from multi-beam sonar. Remote Sens. 4, 3427–3443.

Hass, H.C., Mielck, F., Fiorentino, D., Papenmeier, S., Holler, P., Bartholomä, A., 2017. Seafloor monitoring west of Helgoland (German Bight, North Sea) using the acoustic ground discrimination system RoxAnn. Geo-Mar Lett 37, 125–136.

Hawley, N., 2004. A Comparison of Suspended Sediment Concentrations Measured by Acoustic and Optical Sensors. J. Gt. Lakes Res. 30, 301–309.

Held, P., Schneider von Deimling, J., 2019. New Feature Classes for Acoustic Habitat Mapping—A Multibeam Echosounder Point Cloud Analysis for Mapping Submerged Aquatic Vegetation (SAV). Geosciences 9, 235.

Hellequin, L., Boucher, J.-M., Lurton, X., 2003. Processing of high-frequency multibeam echo sounder data for seafloor characterization. IEEE J. Ocean. Eng. 28, 78–89.

Herkül, K., Peterson, A., Paekivi, S., 2017. Applying multibeam sonar and mathematical modeling for mapping seabed substrate and biota of offshore shallows. Estuar. Coastal. Shelf. Sci 192, 57–71.

Hewitt, J.E., Thrush, S.F., Halliday, J., Duffy, C., 2005. The importance of small-scale habitat structure for maintaining beta diversity. Ecology 86, 1619–1626.

Hewitt, J.E., Thrush, S.F., Legendre, P., Funnell, G.A., Ellis, J., Morrison, M., 2004. Mapping of marine soft-sediment communities: integrated sampling for ecological interpretation. Ecol. Appl. 14, 1203–1216.

Hijmans, R.J., Van Etten, J., Cheng, J., Mattiuzzi, M., Sumner, M., Greenberg, J.A., 2014. Raster: geographic data analysis and modeling. R package version 2.1-16. Vienna, Austria: R Foundation for Statistical Computing.

Hitt, S., Pittman, S.J., Nemeth, R.S., 2011. Diel movements of fishes linked to benthic seascape structure in a Caribbean coral reef ecosystem. Mar. Ecol. Prog. Ser. 427, 275–291.

Hoitink, A.J.F., Hoekstra, P., 2005. Observations of suspended sediment from ADCP and OBS measurements in a mud-dominated environment. Coast. Eng. 52, 103–118.

Holler, P., Markert, E., Bartholomä, A., Capperucci, R., Hass, H.C., Kröncke, I., Mielck, F., Reimers, H.C., 2017. Tools to evaluate seafloor integrity: comparison of multi-device acoustic seafloor classifications for benthic macrofauna-driven patterns in the German Bight, southern North Sea. Geo-Mar. Lett. 37, 93–109.

Holliday, D.V., Pieper, R.E., 1980. Volume scattering strengths and zooplankton distributions at acoustic frequencies between 0.5 and 3 MHz. J. Acoust. Soc. Am. 67, 135–146.

Horn, B.K.P., 1981. Hill shading and the reflectance map. Proc. IEEE 69, 14–47.

Houziaux J.S., Degrendele K., Norro A., Mallefet J., Kerckhof F., Roche M. 2007. Gravel fields of the Western Belgian border, southern bight of the North Sea: a multidisciplinary approach to habitat characterization and mapping. Proceedings of the Conference " UAM2007 -Underwater Acoustic Measurements: Technologies and Results ", Heraklion, Crete, June 2007.

Houziaux, J.-S., Fettweis, M., Francken, F., Van Lancker, V., 2011. Historic (1900) seafloor composition in the Belgian–Dutch part of the North Sea: a reconstruction based on calibrated visual sediment descriptions. Cont. Shelf Res. 31, 1043–1056.

Houziaux, J.S., Kerckhof, F., Degrendele, K., Roche, M., Norro, A., 2008. The Hinder banks: yet an important region for the Belgian marine biodiversity. Bruss. Belg. Sci. Policy 249.

Howarth, T.R., 2015. The Submarine Signal Company. J. Acoust. Soc. Am. 137, 2273–2273.

Huff, L.C., 2008. Acoustic Remote Sensing as a Tool for Habitat Mapping in Alaska Waters. In Marine Habitat Mapping Technology for Alaska; Reynolds, J.R., Greene, H.G., Eds.; Alaska Sea Grant, University of Alaska Fairbanks: Fairbanks, AK, USA; 29–46.

Hughes-Clarke, J., Danforth, B.W., Valentine, P., 1997. Areal seabed classification using backscatter angular response at 95 kHz. In: Pace, N.G., Pouliquen, E., Bergen, O., Lyons, A.P., editors, SACLANTCEN conference proceeding CP-45, Lerici, 1997,243–50

Hughes-Clarke, J.E., Iwanowska, K.K., Parrott, R., Duffy, G., Lamplugh, M., Griffin, J., 2008. Inter-calibrating multi-source, multi-platform backscatter data sets to assist in compiling regional sediment type maps: Bay of Fundy. In: Proceedings of Canadian Hydrographic Conference and National Surveyors Conference.

Hughes-Clarke, J.E., Mayer, L.A., Wells, D.E., 1996. Shallow-water imaging multibeam sonars: A new tool for investigating seafloor processes in the coastal zone and on the continental shelf. Mar. Geophys. Res. 18, 607–629.

Hughes-Clarke, J.H., 1994. Toward remote seafloor classification using the angular response of acoustic backscattering: a case study from multiple overlapping GLORIA data. IEEE J. Oceanic Eng.19, 112–127.

Hussain, M., Chen, D., Cheng, A., Wei, H., Stanley, D., 2013. Change detection from remotely sensed images: From pixel-based to object-based approaches. ISPRS J. Photogramm. Remote Sens. 80, 91–106.

ICES, 2019. Workshop on scoping of physical pressure layers causing loss of benthic habitats D6C1– methods to operational data products (WKBEDLOSS). ICES

Scientific Reports. 1:15. 37. URL http://doi.org/10.17895/ices.pub.5138, (accessed 30.05.2019).

ICES. 2016. Interim Report of the Working Group on Marine Habitat Mapping (WGMHM), 9–11 May 2016, Winchester, UK. ICES CM 2016/SSGEPI:19. 71.

Ierodiaconou, D., Laurenson, L., Leblanc, M., Stagnitti, F., Duff, G., Salzman, S., Versace, V., 2005. The consequences of land use change on nutrient exports: a regional scale assessment in south-west Victoria, Australia. J. Environ. Manage. 74, 305–316.

Ierodiaconou, D., Monk, J., Rattray, A., Laurenson, L., Versace, V.L., 2011. Comparison of automated classification techniques for predicting benthic biological communities using hydroacoustics and video observations. Cont. Shelf Res. 31, S28–S38.

Ierodiaconou, D., Schimel, A.C.G., Kennedy, D., Monk, J., Gaylard, G., Young, M., Diesing, M., Rattray, A., 2018. Combining pixel and object-based image analysis of ultra-high resolution multibeam bathymetry and backscatter for habitat mapping in shallow marine waters. Mar. Geophys. Res. 39, 271–288.

IHO., (5th ed.) 2008. International Hydrographic Organisation Standards for Hydrographic Surveys, February 2008. URL https://www.iho.int/iho_pubs/standard/S-44_5E.pdf, (accessed 24.03.2019).

Innangi, S., Barra, M., Di Martino, G., Parnum, I.M., Tonielli, R., Mazzola, S., 2015. Reson SeaBat 8125 backscatter data as a tool for seabed characterization (Central Mediterranean, Southern Italy): Results from different processing approaches. Appl. Acoust. 87, 109–122.

Ivakin, A.N., 2008. Scattering from inclusions in Marine Sediments: SAX04 Data/Model Comparisons: Defense Technical Information Center, Fort Belvoir, VA. URL https://doi.org/10.21236/ADA494558, (accessed 05.03.2018).

Ivakin, A.N., Sessarego, J.-P., 2007. High frequency broad band scattering from water-saturated granular sediments: Scaling effects. The J. Acoust. Soc. Am. 122, EL165–EL171.

J

Jackson, D.R., Baird, A.M., Crisp, J.J., Thomson, P.A.G., 1986. High-frequency bottom backscatter measurements in shallow water. J. Acoust. Soc. Am. 80, 1188–1199.

James, G., Witten, D., Hastie, T., Tibshirani, R., 2013. An introduction to statistical learning. Springer.

Janowski, L., Trzcinska, K., Tegowski, J., Kruss, A., Rucinska-Zjadacz, M., Pocwiardowski, P., 2018. Nearshore Benthic Habitat Mapping Based on Multi-Frequency, Multibeam Echosounder Data Using a Combined Object-Based Approach: A Case Study from the Rowy Site in the Southern Baltic Sea. Remote Sens. 10, 1983.

Jenkins, C., Eggleton, J., Barry, J., O'Connor, J., 2018. Advances in assessing Sabellaria spinulosa reefs for ongoing monitoring. Ecol. Evol. 8, 7673–7687.

Jenks, G.F., 1967. The data model concept in statistical mapping. Int. Yearb. Cartogr. 7, 186–190.

Johnson, R.K., 1977. Sound scattering from a fluid sphere revisited. J. Acoust. Soc. Am. 61, 375–377.

Jones, D.O.B., Gates, A.R., Huvenne, V.A.I., Phillips, A.B., Bett, B.J., 2019. Autonomous marine environmental monitoring: Application in decommissioned oil fields. Sci. Tot. Environ. 668, 835–853.

Jones, J.B., 1992. Environmental impact of trawling on the seabed: a review. N. Z. J. Mar. Freshw. Res. 26, 59–67.

K

Kaskela, A., Kotilainen, A., Alanen, U., Cooper, R., Green, S., Guinan, J., van Heteren, S., Kihlman, S., Van Lancker, V., Stevenson, A., the EMODnet Geology Partner., 2019. Picking Up the Pieces—Harmonising and Collating Seabed Substrate Data for European Maritime Areas. Geosciences 9, 84.

Kenny, A., 2003. An overview of seabed-mapping technologies in the context of marine habitat classification. ICES J. Mar. Sci. 60, 411–418.

Kerckhof, F., Houziaux, J.S., 2003. Biodiversity of the Belgian marine areas. Biodivers. Belg. K. Belg. Intistuut Voor Natuurwetenschappen Brussel. URL http://www.vliz.be/imisdocs/publications/213322.pdf, (accessed 30.05.2019).

Kimball, P., Rock, S., 2011. Sonar-based iceberg-relative navigation for autonomous underwater vehicles. Deep Sea Research Part II: Topical Studies in Oceanography 58, 1301–1310.

Kint, L., **Montereale-Gavazzi**, G., Van Lancker, V. 2018. Kaderrichtlijn Mariene Strategie. Beschrijvend element 6: Zeebodemintegriteit. Ruimtelijke analyse fysisch verlies en fysische verstoring. Koninklijk Belgisch Instituut voor Natuurwetenschappen: Brussel. 41. URL http://www.vliz.be/imisdocs/publications/313267.pdf, (accessed 30.05.2019).

Klapholz, J., 1988. The History of Sound Reinforcement, in: Audio Engineering Society Conference: 6th International Conference: Sound Reinforcement. Audio Engineering Society. AS International Conference, Nashville, Tennessee, 1988 May 5-8.

Kongsberg Maritime. EM Series Multibeam echo sounders Datagram formats - 850-160692/W. January 2018. URL https://www.km.kongsberg.com/ks/web/nokbg0397.nsf/AllWeb/253E4C58DB98DDA4C1256D790048373B/$file/160692_em_datagram_formats.pdf, (accessed 07.09.2018).

Kongsberg Maritime. Seafloor Information System Reference Manual - 429004/A; 5th ed. 2018. URL https://www.km.kongsberg.com/ks/web/nokbg0397.nsf/AllWeb/A269870356C3A572

C1256E19004ECFA9/$file/164878ac_SIS_Product_specification.pdf, (accessed 07.09.2018).

Korakas, A., Sturm, F., Sessarego, J., Ferrand, D., 2008. Tank experiments of sound propagation over a tilted bottom: Comparison with a 3-D PE model. The J. Acoust. Soc. Am. 123, 3598–3598.

Kostylev, V.E., Todd, B.J., Fader, G.B., Courtney, R.C., Cameron, G.D., Pickrill, R.A., 2001. Benthic habitat mapping on the Scotian Shelf based on multibeam bathymetry, surficial geology and sea floor photographs. Mar. Ecol. Prog. Ser. 219, 121–137.

Kruss, A., Blondel, P., Tegowski, J., 2011. Mapping macrophytes and habitats: single-beam and multibeam imaging in the Arctic, in: 4th International Conference and Exhibition on "Underwater Acoustic Measurements: Technologies and Results. 1687–1694.

Kruss, A., Blondel, P., Tegowski, J., Wiktor, J., Tatarek, A., 2008. Estimation of macrophytes using single-beam and multibeam echosounding for environmental monitoring of Arctic fjords (Kongsfjord, West Svalbard Island). J. Acoust. Soc. Am. 123, 3213–3213.

Kruss, A., Madricardo, F., Sigovini, M., Christian, F., **Montereale-Gavazzi**, G., 2015. Assessment of submerged aquatic vegetation abundance using multibeam sonar in very shallow and dynamic environment. The Lagoon of Venice (Italy) case study, in: 2015 IEEE/OES Acoustics in Underwater Geosciences Symposium (RIO Acoustics). Presented at the 2015 IEEE/OES Acoustics in Underwater Geosciences Symposium (RIO Acoustics), IEEE, Rio de Janeiro, Brazil, 1–7.

Kruss, A., Tęgowski, J., Tatarek, A., Wiktor, J., Blondel, P., 2017. Spatial distribution of macroalgae along the shores of Kongsfjorden (West Spitsbergen) using acoustic imaging. Polish Polar Research 38, 205–229.

Kuhn, M., 2008. Building predictive models in R using the caret package. J. Stat. Softw. 28, 1–26.

Kursa, M.B., Rudnicki, W.R., 2010. Feature selection with the Boruta package. J. Stat. Softw. 36, 1–13.

Kutser, T., Vahtmäe, E., Martin, G., 2006. Assessing suitability of multispectral satellites for mapping benthic macroalgal cover in turbid coastal waters by means of model simulations. Estuar. Coastal. Shelf. Sci 67, 521–529.

L

Lacharité, M., Brown, C.J., 2019. Utilizing benthic habitat maps to inform biodiversity monitoring in marine protected areas. Aquatic Conserv. Mar. Freshw. Ecosyst. 3074.

Lacharité, M., Brown, C.J., Gazzola, V., 2018. Multisource multibeam backscatter data: developing a strategy for the production of benthic habitat maps using semi-automated seafloor classification methods. Mar. Geophys. Res. 39, 307–322.

Ladroit, Y., Lamarche, G., Pallentin, A., 2018. Seafloor multibeam backscatter calibration experiment: comparing 45°-tilted 38-kHz split-beam echosounder and 30-kHz multibeam data. Mar. Geophys. Res. 39, 41–53.

Lamarche, G., Lurton, X., 2018. Recommendations for improved and coherent acquisition and processing of backscatter data from seafloor-mapping sonars. Mar. Geophys. Res. 39, 5–22.

Lamarche, G., Lurton, X., Verdier, A.-L., Augustin, J.-M., 2011. Quantitative characterisation of seafloor substrate and bedforms using advanced processing of multibeam backscatter—Application to Cook Strait, New Zealand. Cont. Shelf Res. 31, S93–S109.

Lanckneus, J., Van Lancker, V., Moerkerke, G., Van Den Eynde, D., Fettweis, M., De Batist, M., Jacobs, P., 2001. Investigation of the natural sand transport on the Belgian continental shelf (BUDGET). Bruss. Fed. Off. Sci. Tech. Cult. Aff. OSTC. URL http://www.belspo.be/belspo/organisation/Publ/pub_ostc/Mn/Abstract/Abs17.pdf, (accessed 30.05.2019).

Lanzoni, J.C., Weber, T.C., 2010. High-resolution calibration of a multibeam echo sounder, in: OCEANS 2010 MTS/IEEE SEATTLE. IEEE, 1–7.

Lasky, M., 1977. Review of undersea acoustics to 1950. J. Acoust. Soc. Am. 61, 283–297.

Lathrop, R.G., Cole, M., Senyk, N., Butman, B., 2006. Seafloor habitat mapping of the New York Bight incorporating sidescan sonar data. Estuar. Coast. Shelf Sci. 68, 221–230.

Lauwaert B., De Witte B., Devriese L., Fettweis M., Martens C., Timmermans S., Van Hoey G., Vanlede J. 2016. Synthesis report on the effects of dredged material dumping on the marine environment (licensing period 2012-2016). RBINS-ILVO-AMT-AMCS-FHR report BL/2016/09. URL http://www.vliz.be/imisdocs/publications/56/229256.pdf, (accessed 30.05.2019).

Le Bot, S., Van Lancker, V., Deleu, S., De Batist, M., Henriet, J.P., 2003. Tertiary and quaternary geology of the Belgian Continental Shelf. Scientific Support Plan for a Sustainable Development Policy. SPSD II North Sea. Brussels, PPS Science policy publication D. 12-75. URL http://www.belspo.be/belspo/organisation/publ/pub_ostc/CPen/CP21Valo_en.pdf, (accessed 10.03.2017).

Lecours, V., Dolan, M.F., Micallef, A., Lucieer, V.L., 2016a. A review of marine geomorphometry, the quantitative study of the seafloor. Hydrol. Earth Syst. Sci. 20.

Levitus, S., Boyer, T.P., 1994. “Salinity and Temperature”, Vol.3 and 4, Salinity and Temperature. World Ocean Atlas, U.S. Department of Commerce National Oceanic and Atmospheric Administration, National Environmental Satellite, Data, and Information Service, Washington D.C. April 1944

Li, J., Tran, M., Siwabessy, J., 2016. Selecting optimal random forest predictive models: a case study on predicting the spatial distribution of seabed hardness. PloS One 11, e0149089.

Liaw, A., Wiener, M., 2002. Classification and regression by randomForest. R News 2, 18–22.

Liu, C., Frazier, P., Kumar, L., 2007. Comparative assessment of the measures of thematic classification accuracy. Remote Sens. Environ. 107, 606–616.

Long, R.D., Charles, A., Stephenson, R.L., 2015. Key principles of marine ecosystem-based management. Mar. Pol. 57, 53–60.

Lu, Y., Qiao, J., Wang, X., 2017. K-normal: An Improved K-means for Dealing with Clusters of Different Sizes, in: Huang, D.-S., Hussain, A., Han, K., Gromiha, M.M. (Eds.), Intelligent Computing Methodologies. Springer International Publishing, Cham, 335–344.

Lucieer, V., Hill, N.A., Barrett, N.S., Nichol, S., 2013. Do marine substrates 'look'and 'sound'the same? Supervised classification of multibeam acoustic data using autonomous underwater vehicle images. Estuar. Coast. Shelf Sci. 117, 94–106.

Lucieer, V., Lamarche, G., 2011. Unsupervised fuzzy classification and object-based image analysis of multibeam data to map deep water substrates, Cook Strait, New Zealand. Cont. Shelf Res. 31, 1236–1247.

Lucieer, V., Lucieer, A., 2009. Fuzzy clustering for seafloor classification. Mar. Geol. 264, 230–241.

Lucieer, V., Picard, K., Siwabessy, J., Jordan, A., Tran, M., Monk, J., 2018. Seafloor Mapping Field Manual for Multibeam Sonar. URL https://www.oceanbestpractices.net/handle/11329/455, (accessed 02.05.2019).

Lucieer, V., Roche, M., Degrendele, K., Malik, M., Dolan, M., Lamarche, G., 2018. User expectations for multibeam echo sounders backscatter strength data-looking back into the future. Mar. Geophys. Res. 39, 23–40.

Lundblad, E., Wright, D.J., Miller, J., Larkin, E.M., Rinehart, R., Battista, T., Anderson, S.M., Naar, D.F., Donahue, B.T., 2006. Classifying benthic terrains with multibeam bathymetry, bathymetric position and rugosity: Tutuila, American Samoa. Mar. Geod 29, 89–111.

Lurton, X., Eleftherakis, D., Augustin, J.-M., 2018. Analysis of seafloor backscatter strength dependence on the survey azimuth using multibeam echosounder data. Mar. Geophys. Res. 39, 183–203.

Lurton, X., Lamarche, G., 2015. Backscatter measurements by seafloor-mapping sonars. Guidelines and Recommendations. URL http://geohab.org/wp-content/uploads/2014/05/BSWGREPORT-MAY2015.pdf, (accessed 07.02.2016).

Lurton, X., 2010. An Introduction to Underwater Acoustics. Principles and Applications, 3rd ed. Springer Verlag Praxis Books and Praxis Publishing, Berlin Heidelnerg.

Lurton, X., 2005. Theoretical Modelling of Acoustical Measurement Accuracy for Swath Bathymetric Sonars. Int. Hydro. Rev. 4, 17-30.

Luyten, P.J., Jones, J.E., Proctor, R., 2003. A numerical study of the long-and short-term temperature variability and thermal circulation in the North Sea. J. Phys. Oceanogr. 33, 37–56.

M

Madricardo, F., Foglini, F., Campiani, E., Grande, V., Catenacci, E., Petrizzo, A., Kruss, A., Toso, C., Trincardi, F., 2019. Assessing the human footprint on the sea-floor of coastal systems: the case of the Venice Lagoon, Italy. Sci. Rep. 9, 6615.

Madricardo, F., Foglini, F., Kruss, A., Ferrarin, C., Pizzeghello, N.M., Murri, C., Rossi, M., Bajo, M., Bellafiore, D., Campiani, E., Fogarin, S., Grande, V., Janowski, L., Keppel, E., Leidi, E., Lorenzetti, G., Maicu, F., Maselli, V., Mercorella, A., **Montereale-Gavazzi**, G., Minuzzo, T., Pellegrini, C., Petrizzo, A., Prampolini, M., Remia, A., Rizzetto, F., Rovere, M., Sarretta, A., Sigovini, M., Sinapi, L., Umgiesser, G., Trincardi, F., 2017. High resolution multibeam and hydrodynamic datasets of tidal channels and inlets of the Venice Lagoon. Sci. Data 4, 170121.

Malik, M., 2019. Sources and Impacts of Bottom Slope Uncertainty on Estimation of Seafloor Backscatter from Swath Sonars. Geosciences 9, 183.

Malik, M., Lurton, X., Mayer, L., 2018. A framework to quantify uncertainties of seafloor backscatter from swath mapping echosounders. Mar. Geophys. Res. 39, 151–168.

Malik, M., Roche, M., Deunf, J.L., Masetti, G., Schimel, A.C.G., Dolan, M.F.J., 2018. Requesting and Comparing Intermediate Backscatter Processing Results from Backscatter Processing Software. Unpublished.

Malvern Panalytical. 2019. Mastersizer 3000. URL https://www.malvernpanalytical.com/en/products/product-range/mastersizer-range/mastersizer-3000/index.html, (accessed 01.03.2019).

Manbachi, A., Cobbold, R.S., 2011. Development and application of piezoelectric materials for ultrasound generation and detection. Ultrasound 19, 187–196.

Marsh, I., Brown, C., 2009. Neural network classification of multibeam backscatter and bathymetry data from Stanton Bank (Area IV). Appl. Acoust. 70, 1269–1276.

Masselink, G., Austin, M.J., O'Hare, T.J., Russell, P.E., 2007. Geometry and dynamics of wave ripples in the nearshore zone of a coarse sandy beach. J. Geophys. Res. 112.

Mathys, M., 2009. The quaternary geological evolution of the Belgian Continental Shelf, southern North Sea. (PhD), Ghent University.

Mathys, M., Van Lancker, V., De Backer, A., Hostens, K., Degrendele, K., Roche, M., 2011. Application for a sand extraction concession in exploration Zone 4: baseline studies on the Hinderbanks and future impact monitoring. In: Study day: marine aggregate extraction: needs, guidelines and future prospects. FOD Economie, KMO, Middenstand en Energie, Bredene. URL http://www.vliz.be/imisdocs/publications/234541.pdf, (accessed 20.11.2016).

Mayer, L., Jakobsson, M., Allen, G., Dorschel, B., Falconer, R., Ferrini, V., Lamarche, G., Snaith, H., Weatherall, P., 2018. The Nippon Foundation—GEBCO Seabed 2030 Project: The Quest to See the World's Oceans Completely Mapped by 2030. Geosciences 8, 63.

Mayer, L.A., 2006. Frontiers in Seafloor Mapping and Visualization. Mar. Geophys. Res. 27, 7–17.

McArthur, M.A., Brooke, B.P., Przeslawski, R., Ryan, D.A., Lucieer, V.L., Nichol, S., McCallum, A.W., Mellin, C., Cresswell, I.D., Radke, L.C., 2010. On the use of abiotic surrogates to describe marine benthic biodiversity. Estuar. Coast. Shelf Sci. 88, 21–32.

McGonigle, C., Grabowski, J.H., Brown, C.J., Weber, T.C., Quinn, R., 2011. Detection of deep water benthic macroalgae using image-based classification techniques on multibeam backscatter at Cashes Ledge, Gulf of Maine, USA. Estuar. Coastal. Shelf. Sci 91, 87–101.

Micallef, A., Le Bas, T.P., Huvenne, V.A., Blondel, P., Hühnerbach, V., Deidun, A., 2012. A multi-method approach for benthic habitat mapping of shallow coastal areas with high-resolution multibeam data. Cont. Shelf Res. 39, 14–26.

Michaelis, R., Hass, H.C., Mielck, F., Papenmeier, S., Sander, L., Ebbe, B., Gutow, L., Wiltshire, K.H., 2019. Hard-substrate habitats in the German Bight (South-Eastern North Sea) observed using drift videos. J. Sea Res. 144, 78–84.

Millard, K., Richardson, M., 2015. On the Importance of Training Data Sample Selection in Random Forest Image Classification: A Case Study in Peatland Ecosystem Mapping. Remote Sens. 7, 8489–8515.

Miller, D.C., Muir, C.L., Hauser, O.A., 2002. Detrimental effects of sedimentation on marine benthos: what can be learned from natural processes and rates? Ecol. Eng. 19, 211–232.

Miller, J.E., Hughes-Clarke, J., Paterson, J., 1997. How effectively have you covered your bottom? Oceanograph. Lit. Rev. 6, 44-46.

Milligan, G.W., Cooper, M.C., 1985. An examination of procedures for determining the number of clusters in a data set. Psychometrika 50, 159–179.

Misiuk, B., Lecours, V., Bell, T., 2018. A multiscale approach to mapping seabed sediments. PloS One 13, e0193647.

Mitchell, P.J., Aldridge, J., Diesing, M., 2019. Legacy Data: How Decades of Seabed Sampling can Produce Robust Predictions and Versatile Products. Geosciences 9, 182.

Mohri, M., Rostamizadeh, A., Talwalkar, A., 2012. Foundations of machine learning. Adaptive computation and machine learning series, MIT Press, Cambridge.

Monk, J., Ierodiaconou, D., Versace, V.L., Bellgrove, A., Harvey, E., Rattray, A., Laurenson, L., Quinn, G.P., 2010. Habitat suitability for marine fishes using presence-only modelling and multibeam sonar. Mar. Ecol. Prog. Ser. 420, 157–174.

Montereale-Gavazzi, G., Madricardo, F., Janowski, L., Kruss, A., Blondel, P., Sigovini, M., Foglini, F., 2016. Evaluation of seabed mapping methods for fine-

scale classification of extremely shallow benthic habitats – Application to the Venice Lagoon, Italy. Estuar. Coast. Shelf Sci. 170, 45–60.

Montereale-Gavazzi, G., Roche, M., Degrendele, K., Lurton, X., Terseleer, N., Baeye, M., Francken, F., Van Lancker, V., 2019. Insights into the Short-Term Tidal Variability of Multibeam Backscatter from Field Experiments on Different Seafloor Types. Geosciences 9, 34.

Montereale-Gavazzi, G., Roche, M., Lurton, X., Degrendele, K., Terseleer, N., Van Lancker, V., 2018. Seafloor change detection using multibeam echosounder backscatter: case study on the Belgian part of the North Sea. Mar. Geophys. Res. 39, 229–247.

Muxika, I., Borja, A., Bonne, W., 2005. The suitability of the marine biotic index (AMBI) to new impact sources along European coasts. Ecol. Indic. 5, 19–31.

N

Naudts, L., Greinert, J., Artemov, Y., Beaubien, S.E., Borowski, C., Batist, M.D., 2008. Anomalous sea-floor backscatter patterns in methane venting areas, Dnepr paleo-delta, NW Black Sea. Mar. Geol. 251, 253–267.

Nechad, B., Ruddick, K.G., Park, Y., 2010. Calibration and validation of a generic multi-sensor algorithm for mapping of total suspended matter in turbid waters. Remote Sens. Environ. 114, 854–866.

O

Olsen, E., Fluharty, D., Hoel, A.H. akon, Hostens, K., Maes, F., Pecceu, E., 2014. Integration at the round table: marine spatial planning in multi-stakeholder settings. PloS One 9, e109964.

P

Parnum I.M., 2007. Benthic Habitat Mapping using Multibeam Sonar System. (PhD), Curtin University, Australia.

Patil, Y.S., Vaidya, M.B., 2012. A technical survey on cluster analysis in data mining. Int. J. Emerg. Technol. Adv. Eng. 2, 503–13.

Paull, L., Seto, M., Leonard, J.J., Li, H., 2018. Probabilistic cooperative mobile robot area coverage and its application to autonomous seabed mapping. The Int. J. Robot Res. 37, 21–45.

Peiman, R., 2011. Pre-classification and post-classification change-detection techniques to monitor land-cover and land-use change using multi-temporal Landsat imagery: a case study on Pisa Province in Italy. Int. J. Remote Sens. 32, 4365–4381.

Pike, R.J., 2000. Geomorphometry-diversity in quantitative surface analysis. Prog. Phys. Geogr. 24, 1–20.

Pike, R.J., Evans, I.S., Hengl, T., 2009. Geomorphometry: a brief guide. Dev. Soil Sci. 33, 3–30.

Pittman, S.J., Christensen, J.D., Caldow, C., Menza, C., Monaco, M.E., 2007. Predictive mapping of fish species richness across shallow-water seascapes in the Caribbean. Ecol. Modell. 204, 9–21.

Pittman, S.J., Kneib, R.T., Simenstad, C.A., 2011. Practicing coastal seascape ecology. Mar. Ecol. Prog. Ser. 427, 187–190.

Pittman, S.J., Olds, A.D., 2015. 34\copyright Simon Pittman, NOAA Seascape ecology of fishes on coral reefs. Ecology of fishes on coral reefs. URL http://research.usc.edu.au/vital/access/manager/Repository/usc:15372, (accessed 10.04.2019).

Pontius Jr, R.G., Millones, M., 2011. Death to Kappa: birth of quantity disagreement and allocation disagreement for accuracy assessment. Int. J. Remote Sens. 32, 4407–4429.

Pontius Jr, R.G., Santacruz, A., 2014. Quantity, exchange, and shift components of difference in a square contingency table. Int. J. Remote Sens. 35, 7543–7554.

Pontius, R.G., Santacruz, A., 2016. diffeR: metrics of difference for comparing Pairs of maps. R package version 0.0–4.

Pontius, R.G., Shusas, E., McEachern, M., 2004. Detecting important categorical land changes while accounting for persistence. Agric. Ecosyst. Environ. 101, 251–268.

Porskamp, P., Rattray, A., Young, M., Ierodiaconou, D., 2018. Multiscale and Hierarchical Classification for Benthic Habitat Mapping. Geosciences 8, 119.

Prampolini, M., Blondel, P., Foglini, F., Madricardo, F., 2018. Habitat mapping of the Maltese continental shelf using acoustic textures and bathymetric analyses. Estuar. Coast. Shelf Sci. 207, 483–498.

Preston, J.M., Christney, A.C., Bloomer, S.F., Beaudet, I.L., 2001. Seabed classification of multibeam sonar images, in: MTS/IEEE Oceans 2001. An Ocean Odyssey. Conference Proceedings (IEEE Cat. No. 01CH37295). IEEE, 2616–2623.

Qimera. Bathyletric procesisng software. URL http://www.qps.nl/display/qimera/Home;jsessionid=DC46DBAC82D50C8F37EAFDF66AD21338, (accessed 02.11.2018).

QPS. URL http://www.qps.nl/display/main/home, (accessed 12.10.2018).

R

R Core Team. 2013. R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing.

Rasband, W.S., 2012. ImageJ: Image processing and analysis in Java. Astrophysics Source Code Library. URL https://imagej.nih.gov/ij/, (accessed 16.07.2018).

Rattray, A., Ierodiaconou, D., Monk, J., Laurenson, L.J.B., Kennedy, P., 2014. Quantification of spatial and thematic uncertainty in the application of underwater video for benthic habitat mapping. Mar. Geod. 37, 315–336.

Rattray, A., Ierodiaconou, D., Monk, J., Versace, V., Laurenson, L., 2013. Detecting patterns of change in benthic habitats by acoustic remote sensing. Mar. Ecol. Prog. Ser. 477, 1–13.

Rattray, A., Ierodiaconou, D., Womersley, T., 2015. Wave exposure as a predictor of benthic habitat distribution on high energy temperate reefs. Front. Mar. Sci. 2, 8.

Rees, H.L., International Council for the Exploration of the Sea (Eds.), 2007. Structure and dynamics of the North Sea benthos, ICES cooperative research report. Internat. Council for the Exploration of the Sea, Copenhagen.

Rice, J., Arvanitidis, C., Borja, A., Frid, C., Hiddink, J.G., Krause, J., Lorance, P., Ragnarsson, S.Á., Sköld, M., Trabucco, B., Enserink, L., Norkko, A., 2012. Indicators for Sea-floor Integrity under the European Marine Strategy Framework Directive. Ecol. Indic. 12, 174–184.

Richards, S.D., Heathershaw, A.D., Thorne, P.D., 1996. The effect of suspended particulate matter on sound attenuation in seawater. J. Acoust. Soc. Am. 100, 1447–1450.

Richardson, M. D., Briggs, K. B., Williams, K. L., Lyons, A. P., Jackson, D. R., 2001. Effects of changing roughness on acoustic scattering: (2) Anthropogenic changes. Proc. Inst. Acoust. 23, 383-390.

Rocchini, D., Boyd, D.S., Féret, J.-B., Foody, G.M., He, K.S., Lausch, A., Nagendra, H., Wegmann, M., Pettorelli, N., 2016. Satellite remote sensing to monitor species diversity: potential and pitfalls. Remote Sens. Ecol. Conserv. 2, 25–36.

Roche, M., Degrendele, K., Vandenreyken, H., Schotte, P., 2017. Multi time and space scale monitoring of the sand extraction and its impact on the seabed by coupling EMS data and MBES measurements. Belgian marine sand: a scarce resource? Belgian FPS Economy — Study day 9. June 2017. URL https://core.ac.uk/download/pdf/84814670.pdf#page=7, (accessed 01.03.2019).

Roche, M., 2002. Utilisation du sonar multifaisceaux pour la classification acoustique des sédiments et son application à la cartographie de la zone de concession 2 de la mer territoriale et du Plateau Continental Belge: etude de faisabilité.

Roche, M., Degrendele, K. Unpublished. The importance of a good calibration of the BS data: pragmatic solutions from IFREMER. [PowerPoint presentation]

Roche, M., Baeye, M., Bisschop, J.D., Degrendele, K., Papili, S., Lopera, O., Lancker, V.V., 2015. Backscatter stability and influence of water column conditions: estimation by multibeam echosounder and repeated oceanographic measurements, belgian part of the North Sea. Proc. Inst. Acoust. 37, 8.

Roche, M., Degrendele, K., Vandenreyken, H., Schotte, P., 2017. Multi time and space scale monitoring of the sand extraction and its impact on the seabed by coupling EMS data and MBES measurements. Belgian marine sand: a scarce resource? Belgian FPS Economy—Study day 9. June 2017, 5–37. URL http://economie.fgov.be/fr/binaries/Articles-study-day-2017_tcm326-283850.pdf, (accessed 01.03.2019).

Roche, M., Degrendele, K., Vrignaud, C., Loyer, S., Le Bas, T., Augustin, J.-M., Lurton, X., 2018. Control of the repeatability of high frequency multibeam echosounder backscatter by using natural reference areas. Mar. Geophys. Res. 39, 89–104.

Roche, M., Le Bas, T.P., Lurton, X., Degrendele, K., De Mol, L., Van Lancker, V., Baeye, M., De Bisschop, J., Vrigneaud, C., Papili, S., Lopera, O., Augustin, J.M., Le Bouffant, N., Berger, L., 2016. Backscatter patch test—inter-comparison of systems using shared reference areas for testing, calibration, and quality assessment. Geological and Biological Marine Habitat Mapping Conference, 15th International Symposium, Salvador Bahia, Brazil. URL http://www.geohab2016.org/sites/geohab16/files/documents/pdf/GeoHab%20 abstract%20booklet.pdf, (accessed 10.10.2016).

Rooper, C.N., Zimmermann, M., 2007. A bottom-up methodology for integrating underwater video and acoustic mapping for seafloor substrate classification. Cont. Shelf Res. 27, 947–957.

Rousseeuw, P.J., 1987. Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. J. Comput. Appl. Math. 20, 53–65.

Rowden, A.A., Jago, C.F., Jones, S.E., 1998. Influence of benthic macrofauna on the geotechnical and geophysical properties of surficial sediment, North Sea. Cont. Shelf Res. 18, 1347–1363.

S

Sandwell, D.T., Smith, W.H., Gille, S., Kappel, E., Jayne, S., Soofi, K., Coakley, B., Géli, L., 2006. Bathymetry from space: Rationale and requirements for a new, high-resolution altimetric mission. Comptes Rendus Geosci. 338, 1049–1062.

Scarrow, R., 2019. Frontiers and deforestation. Nature Plants 5, 124–124.

Schimel, A.C.G., Beaudoin, J., Parnum, I.M., Le Bas, T., Schmidt, V., Keith, G., Ierodiaconou, D., 2018. Multibeam sonar backscatter data processing. Mar. Geophys. Res. 39, 121–137.

SeaBeam Instruments. 2000. L-3 Communications - Multibeam Sonar Theory of Operation. L-3 Communications SeaBeam Instruments. URL https://www3.mbari.org/data/mbsystem/sonarfunction/SeaBeamMultibeamTheoryOperation.pdf, (accessed 01.03.2019).

Shadman Roodposhti, M., Aryal, J., Lucieer, A., Bryan, B., 2019. Uncertainty assessment of Hyperspectral Image Classification: Deep Learning vs. Random Forest. Entropy 21, 78.

Shannon, C.E., 2001. A Mathematical Theory of Communication ACM SIGMOBILE Mob. Comput. Rev. 5, 3-55.

Shaw, G.A., Burke, H.K., 2003. Spectral imaging for remote sensing. Linc. Lab. J. 14, 3–28.

Siemes, K., Snellen, M., Simons, D.G., Hermand, J.-P., Meyer, M., Le Gac, J.-C., 2008. High-frequency multibeam echosounder classification for rapid environmental assessment. J. Acoust. Soc. Am. 123, 3622–3622.

Simard, Y., Stepnowski, A. 2007. Classification methods and criteria. ICES Cooperative Research Report, 286, 64–76.

Simmons, S.M., Parsons, D.R., Best, J.L., Oberg, K.A., Czuba, J.A., Keevil, G.M., 2017. An evaluation of the use of a multibeam echo-sounder for observations of suspended sediment. Appl. Acous. 126, 81–90.

Simmons, S.M., Parsons, D.R., Best, J.L., Orfeo, O., Lane, S.N., Kostaschuk, R., Hardy, R.J., West, G., Malzone, C., Marcus, J., Pocwiardowski, P., 2010. Monitoring Suspended Sediment Dynamics Using MBES. J. Hydr. Eng. 136, 45–49.

Simons, D.G., Snellen, M., 2009. A Bayesian approach to seafloor classification using multi-beam echo-sounder backscatter data. Appl. Acous. 70, 1258–1268.

Singh, A., 1989. Review Article Digital change detection techniques using remotely-sensed data. Int. J. Remote Sens. 10, 989–1003.

Snellen, M., Gaida, T.C., Koop, L., Alevizos, E., Simons, D.G., 2018. Performance of Multibeam Echosounder Backscatter-Based Classification for Monitoring Sediment Distributions Using Multitemporal Large-Scale Ocean Data Sets. IEEE J. Ocean. Eng. 1–14.

Solan, M., Kennedy, R., 2002. Observation and quantification of in situ animal-sediment relations using time-lapse sediment profile imagery (t-SPI). Mar. Ecol. Prog. Ser. 228, 179–191.

SonarScope© - Ifremer Fleet. URL http://flotte.ifremer.fr/fleet/Presentation-of-the-fleet/On-board-software/SonarScope, (accessed 12.10.2018),

Soulsby, R., 1997. Dynamics of marine sands: a manual for practical applications. Thomas Telford, London.

Spearman, J., 2015. A review of the physical impacts of sediment dispersion from aggregate dredging. Mar. Pollut. Bull. 94, 260–277.

Stehman, S.V., Czaplewski, R.L., 1998. Design and analysis for thematic map accuracy assessment: fundamental principles. Remote Sens. Environ. 64, 331–344.

Stephens, D., Diesing, M., 2014. A Comparison of Supervised Classification Methods for the Prediction of Substrate Type Using Multibeam Acoustic and Legacy Grain-Size Data. PLoS ONE 9, e93950.

Stock, A., Crowder, L.B., Halpern, B.S., Micheli, F., 2018. Uncertainty analysis and robust areas of high and low modelled human impact on the global oceans. Conserv. Biol. 32, 1368–1379.

Story, M., Congalton, R.G., 1986. Accuracy assessment: a user's perspective. Photogramm. Eng. Remote Sens. 52, 397–399.

Strong, J.A., 2015. Habitat area as an indicator of Good Environmental Status under the Marine Strategy Framework Directive: the identification of suitable habitat mapping methods with recommendations on best-practice for the reduction of uncertainty.

Strong, J.A., Clements, A., Lillis, H., Galparsoro, I., Bildstein, T., Pesch, R., Birchenough, H. editor: S., 2018. A review of the influence of marine habitat classification schemes on mapping studies: inherent assumptions, influence on

end products, and suggestions for future developments. ICES J. Mar. Sci. 76, 10–22.

Strong, J.A., Elliott, M., 2017. The value of remote sensing techniques in supporting effective extrapolation across multiple marine spatial scales. Mar. Pollut. Bull. 116, 405–419.

Sutherland, T.F., Galloway, J., Loschiavo, R., Levings, C.D., Hare, R., 2007. Calibration techniques and sampling resolution requirements for groundtruthing multibeam acoustic backscatter (EM3000) and QTC VIEW^{TM} classification technology. Estuar. Coastal. Shelf. Sci 75, 447–458.

T

Taylor, P.D., Wilson, M.A., 2002. A new terminology for marine organisms inhabiting hard substrates. Palaios 17, 522–525.

Tęgowski, J., 2005. Acoustical classification of the bottom sediments in the southern Baltic Sea. Quat. Int. 130, 153–161.

Tekman, M.B., Krumpen, T., Bergmann, M., 2017. Marine litter on deep Arctic seafloor continues to increase and spreads to the North at the HAUSGARTEN observatory. Deep Sea Res. Part I: Oceanographic Research Papers 120, 88–99.

Telfer, T.C., Atkin, H., Corner, R.A., 2009. Review of environmental impact assessment and monitoring in aquaculture in Europe and North America. In: Environmental impact assessment and monitoring in aquaculture. FAO Fish Aquac. Tech. Pap. No. 527. UN FAO, Rome, 285−394.

Thorne, P.D., Hanes, D.M., 2002. A review of acoustic measurement of small-scale sediment processes. Cont. Shelf Res. 22, 603–632.

Thrush, S.F., Dayton, P.K., 2002. Disturbance to marine benthic habitats by trawling and dredging: implications for marine biodiversity. Annu. Rev. Ecol. Syst. 33, 449–473.

Tillin, H.M., Luff, A., Davidson, J.J., Perret, J., Huvenne, V., Bett, B.J., Van Rein, H., 2018. Autonomous Underwater Vehicles for use in marine benthic monitoring (JNCC No. 2), Marine Monitoring Platform Guidelines. JNCC, Peterborough. 33, ISSN 2517-7605

Todd, B.J., 2005. Morphology and composition of submarine barchan dunes on the Scotian Shelf, Canadian Atlantic margin. Geomorphology 67, 487–500.

Todd, B.J., Kostylev, V.E., Fader, G.B.J., Courtney, R.C., Pickrill, R.A., 2000. New approaches to benthic habitat mapping integrating multibeam bathymetry and backscatter, surficial geology and sea floor photographs: a case study from the Scotian Shelf, Atlantic Canada, in: ICES 2000 Annual Science Conference, Bruges, Belgium.

Turner, J.A., Babcock, R.C., Hovey, R., Kendrick, G.A., 2018. Can single classifiers be as useful as model ensembles to produce benthic seabed substratum maps? Estuar. Coast. Shelf Sci. 204, 149–163.

U

Underwood, A.J., 1994. On beyond BACI: sampling designs that might reliably detect environmental disturbances. Ecol. Appl. 4, 3–15.

UNFPA. United Nations Population Fund, 2018. State of World Population Report. URL https://www.unfpa.org/sites/default/files/pub-pdf/UNFPA_PUB_2018_EN_SWP.pdf, (accessed 28.03.2019).

Urgeles, R., Locat, J., Schmitt, T., Clarke, J.E.H., 2002. The July 1996 food deposit in the Saguenay Fjord, Quebec, Canada: implications for sources of spatial and temporal backscatter variations. Mar. Geol. 184, 41-60

Urick, R.J., 1948. The Absorption of Sound in Suspensions of Irregular Particles. J. Acoust. Soc. Am. 20, 283–289.

Urick, R.J., 1967. Principles of underwater sound for engineers. Tata McGraw-Hill.

V

Valle, M., Borja, Á., Chust, G., Galparsoro, I., Garmendia, J.M., 2011. Modelling suitable estuarine habitats for Zostera noltii, using ecological niche factor analysis and bathymetric LiDAR. Estuar. Coastal. Shelf. Sci 94, 144–154.

Van den Eynde, D., 2004. Interpretation of tracer experiments with fine-grained dredging material at the Belgian Continental Shelf by the use of numerical models. J. Mar. Syst. 48, 171–189.

Van den Eynde, D., Giardino, A., Portilla, J., Fettweis, M., Francken, F., and Monbaliu, J., 2010. Modelling the effects of sand extraction, on sediment transport due to tides, on the Kwinte Bank. J. Coast. Res. 101-116.

van Denderen, P.D., Bolam, S.G., Hiddink, J.G., Jennings, S., Kenny, A., Rijnsdorp, A.D., Van Kooten, T., 2015. Similar effects of bottom trawling and natural disturbance on composition and function of benthic communities across habitats. Mar. Ecol. Prog. Ser. 541, 31–43.

Van Der Ben, C., Van Der Ben, D., Van Goethem, J., Daro, M., 1977. Inventaire et zonation des organismes sur les brise-lames de la cote beige. Proj. Mer 7, 313–332.

van der Kooij, J., Mackinson, S., Nicholls, A., Righton, D. 2004. Sandeel detection in the sediment using QTC: searching for a needle in a haystack? ICES CM 2004/T:06. URL http://ices.dk/sites/pub/CM%20Doccuments/2004/T/T0604.pdf, (accessed 30.05.2019).

Van Der Reijden, K.J., Koop, L., O'flynn, S., Garcia, S., Bos, O., Van Sluis, C., Maaholm, D.J., Herman, P.M., Simons, D.G., Olff, H., 2019. Discovery of Sabellaria spinulosa reefs in an intensively fished area of the Dutch Continental Shelf, North Sea. J. Sea Res. 144, 85–94.

van Dijk, T.A.G.P., van Dalfsen, J.A., Van Lancker, V., van Overmeeren, R.A., van Heteren, S., Doornenbal, P.J., 2012. Benthic Habitat Variations over Tidal

Ridges, North Sea, the Netherlands, in: Seafloor Geomorphology as Benthic Habitat. Elsevier, 241–249.

Van Hoey, G., Degraer, S., Vincx, M., 2004. Macrobenthic community structure of soft-bottom sediments at the Belgian Continental Shelf. Estuar. Coast. Shelf Sci. 59, 599–613.

Van Lancker, V., 1999. Sediment and morpho dynamics of a siliciclastic near coastal area, in relation to hydrodynamical and meteorological conditions: Belgian continental shelf. (PhD), Gent University.

Van Lancker, V., 2017. Bedforms as Benthic Habitats: living on the edge, chaos, order and complexity, in: Atlas of Bedforms in the Western Mediterranean. Springer, 195–198.

Van Lancker, V., Baeye, M., 2015. Wave glider monitoring of sediment transport and dredge plumes in a shallow marine sandbank environment. PloS One 10, e0128948.

Van Lancker, V., Baeye, M., Evagelinos, D., Francken, F., **Montereale-Gavazzi**, G., Van den Eynde, D., 2017. MSFD-compliant assessment of the physical effects of marine aggregate extraction in the Hinder Banks, synthesis of the first 5 years, in: Degrendele, K. et al. (Ed.) Belgian marine sand: a scarce resource? Study day, 9 June 2017, Hotel Andromeda, Ostend. 87-104.

Van Lancker, V., Baeye, M., **Montereale-Gavazzi**, G., Van den Eynde, D., 2016. Monitoring of the impact of the extraction of marine aggregates, *in casu* sand, in the zone of the Hinder Banks. Period 1/1–31/12 2015 and synthesis of results 2011–2015. Brussels, RBINS-OD Nature. Report MOZ4-ZAGRI/I/VVL/2016/EN/SR01. URL http://www.vliz.be/imisdocs/publications/268906.pdf, (accessed 18.11.2016).

Van Lancker, V., Du Four, I., Verfaillie, E., Deleu, S., Schelfaut, K., Fettweis, M., Van den Eynde, D., Francken, F., Monbaliu, J., Giardino, A., 2007. Management, research and budgetting of aggregates in shelf seas related to end-users (Marebasse). URL http://www.belspo.be/belspo/organisation/publ/pub_ostc/EV/rEV18Ann_en.pdf , (accessed 03.05.2019).

Van Lancker, V., Kint, L., **Montereale-Gavazzi**, G., 2018. Fysische verstoring en verlies van de zeebodem (D6). In: Belgische Staat. Art. 17 Beoordeling voor de Belgische mariene wateren – Richtlijn 2008/59/EG. Kaderrichtlijn Mariene Strategie. BMM, Federale Overheidsdienst Volksgezondheid, Veiligheid van de Voedselketen en Leefmilieu, Brussel, België. URL https://odnature.naturalsciences.be/msfd/nl/assessments/2018/page-d6, (accessed 03.05.2019).

Van Lancker, V., Moerkerke, G., Du Four, I., Verfaillie, E., Rabaut, M., Degraer, S., 2012. Fine-scale geomorphological mapping of sandbank environments for the prediction of macrobenthic occurrences, Belgian part of the North Sea, in: Seafloor Geomorphology as Benthic Habitat. Elsevier, 251–260.

Van Lancker, V., Verfaillie, E., 2005. Geophysical zonation, in: Towards a Spatial Structure Plan for Sustainable Management of the Sea: Mixed Actions-Final Report: SPSD II (MA/02/006). Belgian Science Policy Office, Belgium.

Van Lancker, V.R., Bonne, W.M., Garel, E., Degrendele, K., Roche, M., Van den Eynde, D., Bellec, V.K., Brière, C., Collins, M.B., Velegrakis, A.F., 2010. Recommendations for the sustainable exploitation of tidal sandbanks. J. Coast. Res. 151–164.

van Leeuwen, S., Tett, P., Mills, D., van der Molen, J., 2015. Stratified and nonstratified areas in the North Sea: Long-term variability and biological and policy implications: NORTH SEA STRATIFICATION REGIMES. J. Geophys. Res. Oceans 120, 4670–4686.

van Leeuwen, S., Tett, P., Mills, D., van der Molen, J., 2015. Stratified and nonstratified areas in the North Sea: Long-term variability and biological and policy implications. J. Geophys. Res. Oceans 120, 4670–4686.

van Rein, H., Brown, C.J., Quinn, R., Breen, J., Schoeman, D., 2011. An evaluation of acoustic seabed classification techniques for marine biotope monitoring over broad-scales (> 1 km2) and meso-scales (10 m2–1 km2). Estuar. Coastal. Shelf. Sci. 93, 336–349.

Vanhellemont, Q., Ruddick, K., 2014. Turbid wakes associated with offshore wind turbines observed with Landsat 8. Remote Sens. Environ. 145, 105–115.

Veenstra, H.J., 1969. Gravels of the southern North Sea. Mar. Geol. 7, 449–464.

Verfaillie, E., 2008. Development and validation of spatial distribution models of marine habitats, in support of the ecological valuation of the seabed. (PhD), Ghent University.

Verfaillie, E., Van Lancker, V., Van Meirvenne, M., 2006. Multivariate geostatistics for the predictive modelling of the surficial sand distribution in shelf seas. Cont. Shelf Res. 26, 2454–2468.

Virtasalo, J.J., Korpinen, S., Kotilainen, A.T., 2018. Assessment of the Influence of Dredge Spoil Dumping on the Seafloor Geological Integrity. Front. Mar. Sci. 5, 131.

W

Ward, T.J., Vanderklift, M.A., Nicholls, A.O., Kenchington, R.A., 1999. Selecting marine reserves using habitats and species assemblages as surrogates for biological diversity. Ecol. Appl. 9, 691–698.

Watling, L., Norse, E.A., 1998. Disturbance of the seabed by mobile fishing gear: a comparison to forest clearcutting. Conserv. Biol. 12, 1180–1197.

Watson, J.E., Dudley, N., Segan, D.B., Hockings, M., 2014. The performance and potential of protected areas. Nature 515, 67.

Weber, T.C., Rice, G., Smith, M., 2018. Toward a standard line for use in multibeam echo sounder calibration. Marine Geophysical Research 39, 75–87.

Wedding, L., Lepczyk, C., Pittman, S., Friedlander, A., Jorgensen, S., 2011. Quantifying seascape structure: extending terrestrial spatial pattern metrics to the marine realm. Mar. Ecol. Prog. Ser. 427, 219–232.

Wegmann, M., Leutner, B., Dech, S., 2016. Remote sensing and GIS for ecologists: using open source software. Pelagic Publishing Ltd.

Wells, D.E., Monahan, D., 2002. IHO S44 standards for hydrographic surveys and the variety of requirements for bathymetric data. Hydrogr. J. 9–16.

Wentworth, C.K., 1922. A scale of grade and class terms for clastic sediments. J. Geol. 30, 377–392.

Wernberg, T., Thomsen, M.S., Tuya, F., Kendrick, G.A., Staehr, P.A., Toohey, B.D., 2010. Decreasing resilience of kelp beds along a latitudinal temperature gradient: potential implications for a warmer future. Ecol. Lett. 13, 685–694.

Williams, K.L., Jackson, D.R., Dajun Tang, Briggs, K.B., Thorsos, E.I., 2009. Acoustic Backscattering from a Sand and a Sand/Mud Environment: Experiments and Data/Model Comparisons. IEEE J. Ocean. Eng. 34, 388–398.

Wilson, M.F., O'Connell, B., Brown, C., Guinan, J.C., Grehan, A.J., 2007. Multiscale terrain analysis of multibeam bathymetry data for habitat mapping on the continental slope. Mar. Geod. 30, 3–35.

Woelfl, A.-C., Snaith, H., Amirebrahimi, S., Devey, C., Dorschel, B., Ferrini, V., Huvenne, V.A., Jakobsson, M., Jencks, J., Johnston, G., 2019. Seafloor Mapping–the challenge of a truly global ocean bathymetry. Frontiers in Marine Science 6, 283.

Woodall, L.C., Sanchez-Vidal, A., Canals, M., Paterson, G.L.J., Coppock, R., Sleight, V., Calafat, A., Rogers, A.D., Narayanaswamy, B.E., Thompson, R.C., 2014. The deep sea is a major sink for microplastic debris. Royal Society Open Science 1, 140317–140317.

Y

Yoklavich, M.M., Greene, H.G., Cailliet, G.M., Sullivan, D.E., Lea, R.N., Love, M.S., 2000. Habitat associations of deep-water rockfishes in a submarine canyon: an example of a natural refuge. Fish. Bull. 98, 625–625.

Z

Zajac, R.N., 1999. Understanding the sea floor landscape in relation to impact assessment and environmental management in coastal marine sediments, in: Biogeochemical Cycling and Sediment Ecology. Springer, 211–227.

Zarayskaya, Y., Dorshow, W., X Wigley, R., X Zwolak, K., Bazhenova, E., Sumiyoshi, M., Sattiabaruth, S., Ketter, T., Bohan, A., Roperez, J., Ryzhov, I., Elsaied, M., Wallace, C., 2019. GEBCO-NF Alumni Team Multibeam and HISAS Bathymetric Data Processing and Delivery Workflow developed for Shell Ocean Discovery XPRIZE competition. Geological and Biological Marine Habitat

Mapping Conference, International Symposium, St. Petersburg, Russia, 13-17 May 2019.

Zhao, T., Lazendic, S., Zhao, Y., **Montereale-Gavazzi**, G., Pizurica, A. 2019. Classification of Multibeam Sonar Image Using the Weyl Transform. IEICE Information and Communication Technology Forum, Bydgoszcz, POLAND (submitted).

Zvoleff, A., 2016. GLCM: calculate textures from grey-level co-occurrence matrices (GLCMs). R Package Version 1.

Zyserman, J.A., Fredsøe, J., 1994. Data Analysis of Bed Concentration of Suspended Sediment. J. Hydraul. Eng. 120, 1021–1042.

Abstract

Multibeam echosounding is indispensable for underwater monitoring, with backscatter and bathymetry data enabling Acoustic Seafloor Classification (ASC) and Change Detection (ACD). ASC is a maturing discipline, whilst ACD has remained virtually unexplored. To further develop techniques for the spatio-temporal quantification of seafloor status and dynamics, state-of-the-art hydroacoustic and ground-truth data were acquired in the Belgian Part of the North Sea and were integrated via automated classification routines. ASC research found variable predictive performance between supervised machine learning and unsupervised clustering classification. 300 kHz backscatter discrimination potential is weaker for heterogenous substrates, constraining the spatial structure and information content of the classification scheme. ACD methodologies were developed allowing the acoustic observation of signals of change and quantified the measurement's sensitivity to environmental cyclicity, advancing the phenomenological and acoustical understanding of the dynamic environment: sources and magnitudes that are paramount for the establishment of ACD in environmental monitoring. Multi-parameter sampling datasets need collecting to fine-tune ASC, better interpret field backscatter measurements, and improve classification schemes. Novel data-types, classifiers and predictors need further investigation, which together with knowledge of the system and emerging technologies, ranging from robotics to ecosystem modelling, paves the way for more innovative monitoring of the marine environment.